PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, C12Q 1/68, C12N 15/85, A01K 67/027, C12N 15/67, A61K 48/00	**A1**	**(11) International Publication Number:** **WO 97/10337** **(43) International Publication Date:** 20 March 1997 (20.03.97)

(21) International Application Number: PCT/US96/13482

(22) International Filing Date: 20 August 1996 (20.08.96)

(30) Priority Data:
60/003,784   15 September 1995 (15.09.95)   US

(71) Applicant: BAYLOR COLLEGE OF MEDICINE [US/US];
One Baylor Plaza, Houston, TX 77381 (US).

(72) Inventors: O'MALLEY, Bert, O.; 639 Ramblewood, Houston,
TX 77079 (US). TSAI, Ming-Jer; 6014 Charlotte, Houston,
TX 77005 (US). TSAI, Sophia, Y.; 6014 Charlotte,
Houston, TX 77005 (US). ONATE, Sergio; 8020 Braesmain
#1605, Houston, TX 77025 (US).

(74) Agents: WARBURG, Richard, J. et al.; Lyon & Lyon, Suite
4700, 633 West Fifth Street, Los Angeles, CA 90071 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY,
CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS,
JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD,
MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD,
SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, ARIPO
patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM,
AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT,
BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC,
NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA,
GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
*Before the expiration of the time limit for amending the
claims and to be republished in the event of the receipt of
amendments.*

---

(54) Title: STEROID RECEPTOR COACTIVATOR COMPOSITIONS AND METHODS OF USE

(57) Abstract

The present invention relates to SRC-1 polypeptides, nucleic acids encoding such polypeptides, cells, tissues and animals containing such nucleic acids, antibodies to such polypeptides, assays utilizing such polypeptides, and methods relating to all of the foregoing. Methods for enhancing and inhibiting transcription by providing SRC-1 polypeptides or fragments thereof to a cell.

1

DESCRIPTION

STEROID RECEPTOR COACTIVATOR

COMPOSITIONS AND METHODS OF USE

Field of the Invention

The present invention relates to the novel protein termed steroid receptor coactivator-one ("SRC-1"), nucleotide sequences encoding SRC-1, as well as various products and methods useful for augmenting or downregulating the activity of one or more steroid receptors.

Background of the Invention

The following description of the background of the invention is provided to aid in understanding the invention but is not admitted to be prior art to the invention.

Transcription is a fundamental biological process whereby an RNA molecule is formed upon a DNA template by complementary base pairing that is mediated by RNA polymerase II. Accumulated evidence indicates that numerous general transcription factors undergo a defined order of assembly at promoter-DNA elements to assure RNA polymerase II binding and initiation of transcription of target genes (for review see Zawel, L. and Reinberg, D. (1995), Ann. Rev. Biochem. 64, 533-561).

2

Activation of transcription can be achieved by direct interaction of activators with one or more components of the basal transcriptional machinery. Direct interaction between activators and the basal
5   transcription machinery has been described for several activators (Stringer, K.F. et al., (1990), _Nature_ 345, 783-785; Kashanchi, F. et al., (1994), _Nature_ 367, 295-299; Sauer, F. et al., (1995), _Nature_ 375, 162-164).

Optimal transactivation by an activator is
10   likely to require additional factors termed adaptors or coactivators. These factors seem to play a key regulatory role in bridging or stabilizing the activator with general transcription factors in the core transcriptional machinery. The ability of an activator
15   to squelch or inhibit transactivation of a target gene by another transactivator suggests that they compete for a limited amount of cofactor(s) required for the transactivation process and furthers the concept that coactivators are required for efficient
20   transactivational function (Flanagan, P.M. et al., (1991), _Nature_ 350, 436-438).

It has been postulated that steroid receptors regulate transcription via interactions with the basal transcriptional machinery. However, the finding that
25   squelching occurs between members of the steroid receptor superfamily, indicates that an additional factor(s) or coactivator(s) is important for efficient ligand-inducible target gene expression by members of this superfamily (Meyer, M.E. et al., (1989), _Cell_ 57,

3

433-442; Conneely, O.M. et al., (1989), "Promoter
specific activating domains of the chicken progesterone
receptor." In Gene Regulation by Steroid Hormones IV.
A.K. Roy and J. Clark, eds. (New York, Berlin,
5   Heidelberg, London, Paris, Tokyo: Springer-Verlag), pp.
220-223; Bocquel, M.T. et al., (1989), Nucleic Acids
Res. 17, 2581-2594; Shemshedini, L. et al., (1992), J.
Biol. Chem. 261, 1834-1839). No such functional
coactivator for this superfamily has previously been
10   identified.

Steroid receptors belong to a superfamily of
ligand inducible transcription factors which regulate
hormone responsive genes and thereby affect several
biological processes including cell growth and
15   differentiation. The steroid/thyroid hormone receptor
superfamily can be divided into two types (termed A and
B) based on their characteristic association with heat
shock proteins, binding to DNA and their ligand-
dependent transactivation function (Tsai, M.-J. and
20   O'Malley, B.W. (1994), Ann. Rev. Biochem. 63, 451-486).

One steroid receptor, the human progesterone
receptor (hPR), is expressed in cells as two isoforms:
PR_B of 120 kDa and PR_A of 94 kDa. The A isoform is a
shorter transcript of PR, lacking the most N-terminal
25   164 amino acids of the B receptor (Kastner, P. et al.,
(1990), EMBO J. 9, 1603-1614; Wei, L.L et al., (1987),
Biochem. 26, 6262-6272). Although they display similar
ligand specificities and DNA-binding affinities in
vitro, the transcriptional activity of the two receptor

4

isoforms show different promoter and cell specificities
when assayed in intact cells (Chalepakis, G. et al.,
(1988), _Cell_ 53, 371-382; Tora, L. et al., (1988),
Nature 333, 185-188; Tung, L. et al., (1993), _Mol._
5   _Endocrinol._ 7, 1256-1265).

Like other members of the steroid receptor
superfamily, the hPRs are modular proteins containing a
ligand binding domain (LBD) at the C-terminus and a
centrally located DNA binding domain (DBD). Two regions
10   in hPR have been thought to contain transcriptional
activation functions (AFs). One is located at the N-
terminus (AF1) and the other (AF2) is located within the
LBD (Tora, L. et al., (1989), _Cell_ 59, 477-487;
Gronemeyer, H. (1991), _Ann. Rev. Genet._ 25, 89-123).
15   Recent results indicate that the hPR$_B$-specific 164 amino
acid fragment may contain an additional activation
function (Sartorius, C.A. et al., (1994), _Mol._
Endocrinol. 8, 1347-1360) that is required for maximal
transactivation of the full-length receptor.
20   Activation of a steroid receptor is a complex
multi-step process that involves structural and
functional alterations of receptor which promote
specific binding to DNA hormone-responsive elements
(HREs) to modulate the target gene expression (for
25   review see Tsai, M.-J. and O'Malley, B.W. (1994), _Ann._
Rev. Biochem. 63, 451-486). Thus, steroid receptors
must undergo a rather complex multi-step activation
process to achieve their ultimate transactivational
function.

5

Coactivators have been implicated widely in
nuclear steroid receptor function. Transcriptional
interference experiments between members of the steroid
receptor superfamily suggested that coactivators are
5   limiting and interact, either directly or indirectly,
with the receptor protein in vivo to modulate
transcription. However, as noted above, no such
functional coactivator for this superfamily has
previously been identified.

10   Summary of the Invention
The present invention relates to SRC-1
polypeptides, nucleic acids encoding such polypeptides,
cells containing such nucleic acids, antibodies to such
gene products, assays utilizing such polypeptides, and
15   methods relating to all of the foregoing. In
particular, this invention relates to methods for
augmenting or downregulating the activity of one or more
steroid receptors.

The present invention is based upon the isola-
20   tion and characterization of a new protein which we have
designated steroid receptor coactivator-1, or SRC-1. We
have determined that modulation of SRC-1 activity is
useful in therapeutic procedures and thus the present
invention provides several agents and methods useful for
25   modulating steroid hormone responses and activities,
including modulation of the activity of other
transactivators.

6

The isolated, purified, and/or enriched SRC-1 polypeptides and/or nucleic acids can be used to transactivate a steroid receptor and thereby promote the level of transcription in an organisms or cell.

5   Administration of the appropriate material can be accomplished by one skilled in the art using methods described herein. For example, one or more transfected and/or transformed cells can be used to perform a gene therapy based treatment where activity of steroid

10   receptors may be involved. Examples of disorders or conditions that involve the activity of steroid receptors include malignancies of the reproductive endocrine system and inflamation and immunity disorders, such as those described in U.S. Patent Application

15   Serial No. 08/479,913, filed June 7, 1995, incorporated herein by reference in its entirety, including any drawings. Examples of other disorders or conditions are listed in references available to those skilled in the art such as the Physicians' Desk Reference and include

20   endocrine disorders, rheumatic disorders, collagen disorders, dermatologic diseases, allergic states, ophthalmic diseases, gastrointestinal diseases, respitory diseases, hematologic disorders, breast cancer, endometriosis, hyperproliferative disorders

25   including cancer and others. Alternatively, methods of the invention may be used to inhibit transcription. For example, a truncated form of SRC-1 can be used as a dominant negative inhibitor of receptor activity.

7

We describe herein the cloning and
characterization of a cDNA encoding a protein required
for hPR transactivational function, hereafter termed
steroid receptor coactivator-one (SRC-1). SRC-1
5   directly and specifically interacts with the ligand
binding domain (LBD) of hPR in a hormone-dependent
manner. Binding of the antagonist RU486 to the receptor
protein abolishes this interaction. Coexpression of
SRC-1 with steroid receptors enhances (>10 fold) the
10   hormone-induced transcription of a cellular target gene
without altering the basal activity. Furthermore,
overexpression of SRC-1 can reverse the ability of ER to
squelch PR-mediated transactivation. Finally,
coexpression of a truncated form of SRC-1, which retains
15   the ability to interact with receptor, results in a
dominant-negative inhibition of receptor activity. SRC-
1 thus encodes a protein that fulfills the properties of
a coactivator which ensures efficient ligand-dependent
activity of steroid receptors on target genes.

20   Thus, in a first aspect the invention features
an isolated, enriched, or purified nucleic acid encoding
a SRC-1 polypeptide.

By "a SRC-1 polypeptide" is meant 25
(preferably 30, more preferably 35, most preferably 40)
25   or more contiguous amino acids set forth in the full
length amino acid sequence of Figure 1, or a functional
derivative thereof as described herein. In certain
aspects, polypeptides of 50, 100, 425, 430, 435, 440 or
more amino acids are preferred. The SRC-1 polypeptide

8

can be encoded by a full-length nucleic acid sequence or
any portion of the full-length nucleic acid sequence, so
long as a functional activity of the polypeptide is
retained. Such functional activity can be, for example,

5   (1) the ability to interact with the PR in an agonist
specific manner, (2) the ability to enhance the hormone-
induced transcriptional activity without altering basal
activity of the promoter, (3) the ability to stimulate
transactiviation of one or all steroid receptors, (4)

10   the ability to reverse ER squelching of hPR activation
in a dose dependent manner; and/or (5) the ability of a
truncated form of the SRC-1 polypeptide to inhibit
receptor activity in a dominant negative manner. The
amino acid sequence is preferably substantially similar

15   to the sequence shown in Figure 1, or fragments thereof.
A sequence that is substantially similar will have at
least 90% identity (preferably at least 95% and most
preferably 99-100%) to the sequence of Figure 1.

By "identity" is meant a property of sequences

20   that measures their similarity or relationship.
Identity is measured by dividing the number of identical
residues by the total number of residues and multiplying
the product by 100. Thus, two copies of exactly the
same sequence have 100% identity, but sequences that are

25   less highly conserved and have deletions, additions, or
replacements may have a lower degree of identity. Those
skilled in the art will recognize that several computer
programs are available for determining sequence
identity.

By "isolated" in reference to nucleic acid is meant a polymer of 2 (preferably 21, more preferably 39, most preferably 75) or more nucleotides conjugated to each other, including DNA or RNA that is isolated from a

5   natural source or that is synthesized. In certain embodiments of the invention longer nucleic acids are preferred, for example those of 1202, 1221, 1239, 1275 or more nucleotides and/or those having at least 50%, 60%, 75%, 90%, 95% or 99% identity to the full length

10 sequence shown in Figure 1. The isolated nucleic acid of the present invention is unique in the sense that it is not found in a pure or separated state in nature. Use of the term "isolated" indicates that a naturally occurring sequence has been removed from its normal

15 cellular environment. Thus, the sequence may be in a cell-free solution or placed in a different cellular environment. The term does not imply that the sequence is the only nucleotide chain present, but that it is essentially free (about 90 - 95% pure at least) of

20 non-nucleotide material naturally associated with it and thus is meant to distinguish from isolated chromosomes.

By the use of the term "enriched" in reference to nucleic acid is meant that the specific DNA or RNA sequence constitutes a significantly higher fraction (2

25 - 5 fold) of the total DNA or RNA present in the cells or solution of interest than in normal or diseased cells or in the cells from which the sequence was taken. This could be caused by a person by preferential reduction in the amount of other DNA or RNA present, or by a

10

preferential increase in the amount of the specific DNA
or RNA sequence, or by a combination of the two.
However, it should be noted that enriched does not imply
that there are no other DNA or RNA sequences present,
5   just that the relative amount of the sequence of
interest has been significantly increased. The term
significant here is used to indicate that the level of
increase is useful to the person making such an
increase, and generally means an increase relative to
10   other nucleic acids of about at least 2 fold, more
preferably at least 5 to 10 fold or even more. The term
also does not imply that there is no DNA or RNA from
other sources. The other source DNA may, for example,
comprise DNA from a yeast or bacterial genome, or a
15   cloning vector such as pUC19. This term distinguishes
from naturally occurring events, such as viral
infection, or tumor type growths, in which the level of
one mRNA may be naturally increased relative to other
species of mRNA. That is, the term is meant to cover
20   only those situations in which a person has intervened
to elevate the proportion of the desired nucleic acid.

        It is also advantageous for some purposes that
a nucleotide sequence be in purified form. The term
"purified" in reference to nucleic acid does not require
25   absolute purity (such as a homogeneous preparation);
instead, it represents an indication that the sequence
is relatively purer than in the natural environment
(compared to the natural level this level should be at
least 2-5 fold greater, e.g., in terms of mg/ml).

11

Individual clones isolated from a cDNA library may be
purified to electrophoretic homogeneity. The claimed
DNA molecules obtained from these clones could be
obtained directly from total DNA or from total RNA. The
5   cDNA clones are not naturally occurring, but rather are
preferably obtained via manipulation of a partially
purified naturally occurring substance (messenger RNA).
The construction of a cDNA library from mRNA involves
the creation of a synthetic substance (cDNA) and pure
10   individual cDNA clones can be isolated from the
synthetic library by clonal selection of the cells
carrying the cDNA library. Thus, the process which
includes the construction of a cDNA library from mRNA
and isolation of distinct cDNA clones yields an
15   approximately 10^6-fold purification of the native
message. Thus, purification of at least one order of
magnitude, preferably two or three orders, and more
preferably four or five orders of magnitude is expressly
contemplated.

20   In preferred embodiments the isolated nucleic
acid comprises, consists essentially of, or consists of
a nucleic acid sequence set forth in the full length
amino acid sequence of Figure 1, a functional derivative
thereof, or encodes at least 75, 90, 105, 120, 150, 475,
25   490, 505, 520, or 550 contiguous amino acids thereof;
the SRC-1 polypeptide comprises, consists essentially
of, or consists of at least 25, 30, 35, or 40 contiguous
amino acids of a SRC-1 polypeptide. The nucleic acid
may be isolated from a natural source by cDNA cloning or

12

subtractive hybridization; the natural source may be
mammalian (human) blood, semen, or tissue and the
nucleic acid may be synthesized by the triester method
or by using an automated DNA synthesizer. In yet other
5   preferred embodiments the nucleic acid is a conserved or
unique region, for example those useful for the design
of hybridization probes to facilitate identification and
cloning of additional polypeptides, the design of PCR
probes to facilitate cloning of additional polypeptides,
10   and obtaining antibodies to polypeptide regions.

By "conserved nucleic acid regions", are meant
regions present on two or more nucleic acids encoding a
SRC-1 polypeptide, to which a particular nucleic acid
sequence can hybridize under lower stringency
15   conditions. Examples of lower stringency conditions
suitable for screening for nucleic acid encoding SRC-1
polypeptides are provided in Abe, et al. _J. Biol. Chem._,
_19_:13361 (1992) (hereby incorporated by reference herein
in its entirety, including any drawings). Preferably,
20 ·   conserved regions differ by no more than 5 out of 20
nucleotides.

By "unique nucleic acid region" is meant a
sequence present in a full length nucleic acid coding
for a SRC-1 polypeptide that is not present in a
25   sequence coding for any other naturally occurring
polypeptide. Such regions preferably comprise 30 or 45 ·
contiguous nucleotides present in the full length
nucleic acid encoding a SRC-1 polypeptide. In

13

particular, a unique nucleic acid region is preferably
of mammalian origin.

The invention also features a nucleic acid
probe for the detection of a SRC-1 polypeptide or
5   nucleic acid encoding a SRC-1 polypeptide in a sample.
The nucleic acid probe contains nucleic acid that will
hybridize to a sequence set forth in Figure 1 or a
functional derivative thereof. The SRC-1 polypeptide
that is detected may comprise, consist of, or consist
10   essentially of any given number of contiguous amino
acids of the amino acid sequence set forth in Fig. 1.

By "comprising" it is meant including, but not
limited to, whatever follows the word "comprising".
Thus, use of the term "comprising" indicates that the
15   listed elements are required or mandatory, but that
other elements are optional and may or may not be pre-
sent. By "consisting of" is meant including, and
limited to, whatever follows the phrase "consisting of".
Thus, the phrase "consisting of" indicates that the
20   listed elements are required or mandatory, and that no
other elements may be present. By "consisting essen-
tially of" is meant including any elements listed after
the phrase, and limited to other elements that do not
interfere with or contribute to the activity or action
25   specified in the disclosure for the listed elements.
Thus, the phrase "consisting essentially of" indicates
that the listed elements are required or mandatory, but
that other elements are optional and may or may not be

14

present depending upon whether or not they affect the activity or action of the listed elements.

In preferred embodiments the nucleic acid probe hybridizes to nucleic acid encoding at least 12,

5   25, 50, 75, 90, 100, 120, 150, 412, 425, 450, 475, 490, 500, 520, or 550 contiguous amino acids of the full-length sequence set forth in Figure 1 or a functional derivative thereof. Various low or high stringency hybridization conditions may be used depending upon the

10   specificity and selectivity desired. Under stringent hybridization conditions only highly complementary nucleic acid sequences hybridize. Preferably, such conditions prevent hybridization of nucleic acids having 1 or 2 mismatches out of 20 contiguous nucleotides.

15   Methods for using the probes include detecting the presence or amount SRC-1 RNA in a sample by contacting the sample with a nucleic acid probe under conditions such that hybridization occurs and detecting the presence or amount of the probe bound to SRC-1 RNA.

20   The nucleic acid duplex formed between the probe and a nucleic acid sequence coding for a SRC-1 polypeptide may be used in the identification of the sequence of the nucleic acid detected (for example see, Nelson *et al.*, in <u>Nonisotopic DNA Probe Techniques</u>, p. 275 Academic

25   Press, San Diego (Kricka, ed., 1992) hereby incorporated by reference herein in its entirety, including any drawings). Kits for performing such methods may be constructed to include a container means having disposed therein a nucleic acid probe.

15

The invention also features recombinant
nucleic acid, preferably in a cell or an organism. The
invention also provides a recombinant cell or tissue
containing a purified nucleic acid coding for a SRC-1

5   polypeptide. The recombinant nucleic acid may contain a
sequence set forth in Figure 1 or a functional
derivative thereof and a vector or a promoter effective
to initiate transcription in a host cell. The
recombinant nucleic acid can alternatively contain a

10   transcriptional initiation region functional in a cell,
a sequence complimentary to an RNA sequence encoding a
SRC-1 polypeptide and a transcriptional termination
region functional in a cell. In such cells, the nucleic
acid may be under the control of its genomic regulatory

15   elements, or may be under the control of exogenous
regulatory elements including an exogenous promoter. By
"exogenous" it is meant a promoter that is not normally
coupled *in vivo* transcriptionally to the coding sequence
for the SRC-1 polypeptide.

20   In other aspects, the invention provides
transgenic, nonhuman mammals containing a transgene
encoding a SRC-1 polypeptide or a gene effecting the
expression of a SRC-1 polypeptide. Such transgenic
nonhuman mammals are particularly useful as an *in vivo*

25   test system for studying the effects of introducing a
SRC-1 polypeptide, regulating the expression of a SRC-1
polypeptide (i.e., through the introduction of
additional genes, antisense nucleic acids, or
ribozymes).

16

A "transgenic animal" is an animal having cells that contain DNA which has been artificially inserted into a cell, which DNA becomes part of the genome of the animal which develops from that cell.

5   Preferred transgenic animals are primates, mice, rats, cows, pigs, horses, goats, sheep, dogs and cats. The transgenic DNA may encode for a human SRC-1 polypeptide. Native expression in an animal may be reduced by providing an amount of anti-sense RNA or DNA effective

10   to reduce expression of the receptor.

In other embodiments a steroid ligand activates a molecular switch (as described herein and in U.S. Patent Application Serial No. 07/939,246, filed September 2, 1992 and 08/479,913, filed June 7, 1995

15   both of which are incorporated herein by reference in their entirety including any drawings) and the SRC-1 polypeptide and thereby provides a super-physiological response to enhance steroid therapy. Expression of the SRC-1 polypeptide may be driven by a constitutively

20   active promoter with a coding region for SRC-1. The gene switch may optionally be provided either in the same plasmid or in a different plasmid. The promoter for SRC-1 may be regulated by a gene switch so that the ligand activates both SRC-1 and the gene of interest.

25   In another aspect the present invention provides a method for increasing the transcription of a target gene. Transcription refers to the process of converting genetic information from DNA to RNA. The method involves the step of providing nucleic acid

17

encoding a SRC-1 polypeptide to a cell containing said

target gene. The increase may be from an initial level

of no transcription or may be from a pre-existing level

of transcription. The target gene can be any gene that

5   is transactivated by SRC-1. The level of transcription

may be determined using methods known in the art; for

example the level of transcription may be assessed by

measuring the chloramphenicol acetyl transferase

activity. Providing SRC-1 nucleic acid, or the

10   polypeptide itself, to a cell can increase the

transcriptional activity of any steroid receptor, such

as the mineral corticoid (MR), androgen (AR), estrogen,

progesterone, Vitamin D, COUP-TF, cis-retonic acid,

Nurr-1, thyroid hormone, mineralocorticoid,

15   glucocorticoid-α, glucocorticoid-β and orphan receptors.

In preferred embodiments the method may also

involve the step of providing a molecular switch for

regulating expression of a nucleic acid cassette in gene

therapy to the cell containing the target gene. The

20   molecular switch includes a natural steroid receptor DNA

binding domain linked to a modified ligand binding

domain. Preferably the SRC-1 polypeptide comprises the

full length amino acid sequence of Fig. 1 or a fragment

thereof having at least 700, 800, or 900 contiguous

25   amino acids of the full length sequence, or a fragment

containing an essential interaction domain of SRC-1.

The switch is preferably tissue specific, as described

herein.

18

The method may also involve: (1) attaching the
molecular switch to a nucleic acid cassette to form a
nucleic acid cassette/molecular switch complex for use
in the gene therapy; (2) administering a pharmacological

5   dose of the nucleic acid cassette/molecular switch
complex to an animal or human to be treated; (3) turning
the molecular switch on or off by dosing the animal or
human with a pharmacological dose of a ligand which
binds to the modified ligand binding site; and (4)

10   transcribing the nucleic acid to produce a protein after
the animal or human is given a pharmacological dose of
the ligand. These steps are described and definitions
for terms such as "nucleic acid cassette", and "plasmid"
are provided in U.S. Patent Application Serial No.

15   07/939,246, filed September 2, 1992 and International
Patent Publication WO 93/23431, published November 25,
1993, both of which are incorporated herein by reference
in their entirety including any drawings.

The molecular switch and the nucleic acid

20   cassette may be on the same or separate plasmids and may
be co-injected into a target cell or injected
separately. Similarly, the molecular switch and the
nucleic acid encoding the SRC-1 polypeptide may be on
the same or separate plasmids and may be co-injected or

25   separately injected into a target cell.

In another aspect the invention provides a
composition of matter comprising a molecular switch
linked to a nucleic acid cassette. The
cassette/molecular switch complex is positionally and

19

sequentially oriented in a vector such that the nucleic
acid in the cassette can be transcribed and when
necessary translated in a target cell. The molecular
switch regulates a constitutively active promoter in a
5   plasmid with a coding region for a SRC-1 polypeptide.

          The invention also features a method for
decreasing the transcription of a target gene. The
method involves providing nucleic acid encoding a
dominant-negative inhibitor of a SRC-1 polypeptide in a
10   cell containing said target gene. The dominant negative
inhibitor preferably is encoded by a N truncated
fragment of the full length sequence, such as the
approximately 150 amino acid long fragment of Example 8.

          In another aspect the present invention
15   provides a molecular switch for regulating expression of
a nucleic acid cassette in gene therapy, comprising a
modified SRC-1 polypeptide, said polypeptide including a
natural SRC-1 activation domain linked to a modified
binding domain. In this embodiment the SRC-1 nucleic
20   acid forms part of the molecular switch. Thus, a
substitution is envisioned in a previosly described
switch which included a DNA binding domain of a steroid
receptor (for example GAL-4) linked to a transactivation
domain (for example VP-16) linked to ligand binding
25   domain (for example a mutated LBD of the progesterone
receptor). The substitution involves replacing VP-16 or
some other transactivation domain with an essential
interaction domain of SRC-1. The term "essential
interaction domain" refers to the portion of SRC-1

20

required for interaction with other transcriptional
factors and agents and those skilled in the art may
locate an essential interaction domain using techniques
known in the art.

5   The invention also features a method for
regulating expression of a nucleic acid cassette in gene
therapy comprising the step of attaching a modified SRC-
1 polypeptide molecular switch to a nucleic acid
cassette to form a nucleic acid/molecular switch complex

10   for use in gene therapy and administering a
pharmacological dose of the nucleic acid
cassette/molecular switch complex to an animal or human
to be treated.

           In another aspect the invention features a

15   composition of matter comprising a modified SRC-1
polypeptide molecular switch linked to a nucleic acid
cassette, wherein said complex is positionally oriented
in a vector such that the nucleic acid in the cassette
can be transcribed and when necessary translated in a

20   target cell.

           The invention also features a method of
treating a SRC-1 related disease or condition (such as
those described herein which require modulation of
steroid receptor activity) comprising the steps of

25   inserting an expression vector containing a SRC-1 coding
sequence into cells, growing the cells in vitro, and
infusing the cells into a patient in need of such
treatment.

21

The summary of the invention described above
is non-limiting and other features and advantages of the
invention will be apparent from the following descrip-
tion of the preferred embodiments, and from the claims.

5   Brief Description of the Drawings
                 Figure 1 shows the nucleotide and amino acid
sequences of SRC-1. The numbers to the left correspond
to nucleotides and amino acid numbers (in brackets). The
underlined methionine at the beginning of amino acid
10   sequence is the putative translation start site. The
regions of the protein rich in S and T (Ñ) and Q (ŸŸŸ)
residues and the stop codon (*) are also indicated.

             Description of the Preferred Embodiments
                 The present invention relates to SRC-1
15   polypeptides, nucleic acids encoding such polypeptides,
cells, tissues and animals containing such nucleic
acids, antibodies to such polypeptides, assays utilizing
such polypeptides, and methods relating to all of the
foregoing.
20   A yeast two-hybrid system was used to identify
proteins that interact with the activation function of
the ligand-binding domain of the human progesterone
receptor (hPR). A protein encoded by a 0.8 kb cDNA from
a B lymphocyte cDNA expression library was found to
25   specifically interact with the receptor in an agonist
specific manner as demonstrated both in the two-hybrid
system and by an in vitro interaction assay using

22

baculovirus expressed PR. The interaction was
significantly lower with ligand-free receptor or
receptor bound by the antagonist RU486.

5   A full-length cDNA which encodes a protein
with calculated mass of 114.1 kDa was isolated.
Coexpression of this cDNA with hPR in mammalian cells
resulted in >10 fold enhancement of the hormone-induced
transcriptional activity, without altering basal
activity of the promoter. In addition, it stimulates
10   transactivation of all the steroid receptors tested,
including GR (glucocorticoid), ER (estrogen), TR, RXR,
and orphan receptors. Therefore, it is termed steroid
receptor coactivator-one, SRC-1. SRC-1 also stimulates
GAL4-VP16 and Sp1 transactivation. In contrast, the
15   transcriptional activity of nuclear factors such as CREB
and E2F was not affected. Furthermore, coexpression of
SRC-1 reverses the ability of the estrogen receptor to
squelch activation by hPR in a dose-dependent manner.
Finally, the N-terminal truncated form of SRC-1 acts as
20   dominant-negative repressor in steroid receptor
transactivation. Taken together, our results indicate
that SRC-1 encodes a coactivator required for full
transactivational activity of the ligand-dependent
steroid receptor superfamily.

25   As noted above, we have successfully used the
yeast two-hybrid system to clone and identify a steroid
receptor coactivator (SRC-1) protein. SRC-1 is
expressed in all human tissues and cell lines analyzed
as two mRNA species. A predominant single mRNA is

23

observed in tissue from brain. With the exception of a glutamine rich region, the predicted open reading frame of 1061 amino acids does not contain an obvious transcriptional regulatory domain. SRC-1 shows higher

5   glutamine (10.6%), serine (12.2%) and leucine (9.2%) residues than do many other proteins. The C-terminal half of SRC-1, which contains the receptor-interacting region, is relatively hydrophobic when compared to the more hydrophilic N-terminal half which has an enriched

10   serine and threonine content.

I.   **Role of SRC-1 in the multi-step steroid receptor transactivation.**

An assay to determine whether a protein acts as a coactivator (Dynlacht, B.D. et al., (1991), Cell

15   66, 563-576; Flanagan, P.M. et al., (1991), Nature 350, 436-438) indicates that SRC-1 is in fact a coactivator. First, coactivators should allow efficient activator-dependent transcription without altering basal activity (Dynlacht, B.D. et al., (1991), Cell 66, 563-576). As

20   shown in this study, overexpression of SRC-1 in mammalian cells results in the enhancement of hPR-dependent transactivation by >10 fold without altering basal activity of the promoter. Second, a coactivator should reverse the squelching phenomenon observed when

25   two transacting factors are present (Flanagan, P.M. et al., (1991), Nature 350, 436-438). Indeed, overexpression of hER resulted in ~19 fold reduction in hPR transactivation. Addition of SRC-1 reversed the

24

squelching effect of hER on hPR transactivation in a

dose-dependent manner, by ~16 fold. Hence, we conclude

that SRC-1 is a genuine coactivator for steroid

receptors. Since SRC-1 is incapable of completely

5   reversing ER squelching of PR, it is possible that

additional factor(s) also participated in concert with

SRC-1 to assure full steroid receptor transactivation of

target genes.

II.   SRC-1 and the role of ligand on

10   steroid receptors transactivation.

SRC-1 appears to act as coactivator for both

type A and B steroid receptors. These results suggest

that SRC-1 acts at a step(s) after receptor binds to

DNA. Consistent with this hypothesis, we observed that

15   when type A receptors are overexpressed in cell lines

that allow partial transactivation in the absence of

hormone, presumably due to the ligand-independent

activation pathway for the type A receptors (Smith, C.L.

et al., (1993), Proc. Natl. Acad. Sci. USA 90, 6120-

20   6124), SRC-1 also increased their ability to

transactivate in a ligand-independent manner. The basal

activity of the reporter again remained unaltered.

Similarly, transactivation of type B receptors such as

TR and RXR that exhibit cell-specific basal activity in

25   the absence of ligand were also affected by SRC-1 in

their ligand free state. These data further suggest

that SRC-1 acts at step(s) after receptors bind to their

cognate HREs. Most likely SRC-1 enhances receptor

25

activity by facilitating receptor interaction with the
basal transcriptional machinery.

The finding that SRC-1 cannot alter the
antagonistic effect of RU486 on PR transactivation is an
5   important observation. This result is consistent with
our observations that SRC-1 does not interact
efficiently with PR in the presence of antagonist, both
in vitro and in intact cells. We and others have
suggested previously that the antagonist RU486 induces a
10   distinct conformational change in the receptor molecule
that impairs the ability of receptor to transactivate,
presumably due to the fact that the altered structure
renders the antagonist-receptor complex unable to
interact with other factors required for receptor-
15   mediated transactivation (El-Ashry, D. et al., (1989),
Mol. Endocrinol. 3, 1545-1558; Vegeto, E. et al.,
(1992), _Cell_ 69, 703-713; Allan, G.F., et al., (1992a),
J. Biol. Chem. 267, 19513-19520; Allan, G.F., et al.,
(1992b), _Proc. Natl. Acad. Sci. USA_ 89, 11750-11754;
20   DeMarzo, A.M. et al., (1992), _Biochem._ 31, 10491-10501).

Our current findings substantiate that the
proper conformational change induced by agonist is
necessary for receptor to interact with its coactivator,
SRC-1. We postulate that the inability of antagonist-
25   bound receptor to interact efficiently with SRC-1 leads
to the biological outcome of hormonal antagonist
administration to intact cells since SRC-1 appears to
recognize and discriminate in favor of the agonist bound

26

receptor complex to establish the transcriptional
initiation.

III. SRC-1 is a coactivator for the
steroid receptor superfamily.

5   The existence of cellular proteins which.
enhance steroid receptor transactivation have been
suggested by both biochemical and genetic approaches
(Meyer, M.E. et al., (1989), Cell 57, 433-442; Bocquel,
M.T. et al., (1989), Nucleic Acids Res. 17, 2581-2594;
10   Conneely, O.M. et al., (1989), "Promoter specific
activating domains of the chicken progesterone
receptor." In Gene Regulation by Steroid Hormones IV.
A.K. Roy and J. Clark, eds. (New York, Berlin,
Heidelberg, London, Paris, Tokyo: Springer-Verlag), pp.
15   220-223; Tasset, D. et al., (1990), Cell 62, 1177-1187;
Shemshedini, L. et al., (1992), J. Biol. Chem. 261,
1834-1839; Berkenstam, A. et al., (1992), Cell 69, 401-
412). Recent studies have identified the yeast proteins
SSN6 (McDonnell, D.P. et al., (1992), Proc. Natl. Acad.
20   Sci. USA 89, 10563-10567), SPT6 (Baniahmad et al.,
1995), (Baniahmad C. et al. (1995), Mol. Endocrinol. 9,
34-43), SIN3 (Nawaz, Z. et al., (1994), Mol. Gen. Genet.
245, 724-733), SNF2/SWI2 (Yoshinaga, S.K. et al.,
(1992), Science 258, 1598-1604; Chiba, H. et al.,
25   (1994), Nucleic Acids Res. 22, 1815-1820; Laurent, B.C.
et al., (1993), Genes and Dev. 7, 583-591) and the human
homolog of hbrm protein (Muchardt, C. and Yaniv, M.
(1993), EMBO J. 12, 4279-4290; Singh, P. et al., (1995),

27

Nature 374, 562-565) as potential regulatory proteins
for ER, PR and GR.

Certain of these factors are thought to
function coordinately with activators to enhance
transcription by affecting chromatin structure and
thereby relieving the repressive effect of chromatin on
gene transcription. Alterations in the nucleosome
structure that affect chromatin also have been
implicated in steroid-hormone action (Archer, T.K. et
al., (1992), _Science_ 255, 1573-1576; Schild, C. et al.,
(1993), _EMBO J._ 12, 423-433; Truss, M. et al., (1995),
EMBO J. 14, 1737-1751). It is possible that these
factors' effect on transactivation of steroid receptors
is of an indirect nature.

Sequence analysis of the hbrm protein has
revealed the presence of a conserved domain between
various transcription factors, the bromodomain, which
appears to be important for protein-protein interactions
(Haynes, S.R. et al., (1992), _Nucleic Acids Res._ 20,
2603). In addition, it also contains a presumptive
helicase domain known to be involved in chromatin
decondensation during DNA replication (Laurent, B.C. et
al., (1993), _Genes and Dev._ 7, 583-591). The ability of
the hbrm protein to upregulate GR transactivation is
dependent upon the coexpression of the Rb protein, a
cell-cycle regulatory protein and again, could involve
changes in chromatin structure (Singh et al., 1995). In
addition, a mouse bromodomain containing protein, TIF-1
(Le Douarin, B. et al., (1995), _EMBO J._ 14, 2020-2033),

28

the human homolog of the adaptor Sug-1, thyroid receptor
interacting protein, Trip-1 (Lee et al., 1995b), as well
as other Trips (Seol, W. et al., (1995), Mol.
Endocrinol. 9, 72-85; Lee et al., 1995a), have been
5   isolated and implicated as modulators and/or mediators
of ligand-dependent steroid receptor transactivation.

Although evidence for ligand-dependent
interaction has been provided for most of these
proteins, no definitive evidence indicating coactivator
10   function has been shown. Other less well characterized
proteins that physically interact with ER, in a ligand-
dependent manner have been described, but no functional
role has been assigned (Halachmi, S. et al., (1994),
Science 264, 1455-1458; Cavailles, V. et al, (1994),
15   Proc. Natl. Acad. Sci. USA 91, 10009-10013). SRC-1
shares no significant homology to the proteins above
mentioned nor does it contain similar functional
canonical domains. Therefore, it is likely that the
mechanism by which SRC-1 alters the rate of
20   transcription will differ.

Although the role of SRC-1 as coactivator for
steroid receptor is documented in this report, we cannot
rule out potential effects on chromatin structure.
Nevertheless, our evidence indicates that SRC-1 acts by
25   direct contact with the receptor protein to modulate its
activity. The ability of SRC-1 to stimulate receptor-
mediated transcription and to reverse squelching between
distinct steroid receptors confirms that these
interactions are relevant and probably occur in vivo.

29

The previous observation that steroid receptors can
stabilize the preinitiation transcription complex (Tsai,
S.Y. et al., (1990), J. Biol. Chem. 265, 17055-17061;
Klein-Hitpass, L. et al., (1990), Cell 60, 247-257) lead
5   us to speculate that SRC-1 may facilitate this process.

        The finding that the N-terminal truncated form
of SRC-1, which contains the receptor-interacting
region, acts in a dominant-negative manner and suggests
that the N-terminal region of SRC-1 is responsible for
10   interaction(s) with the basal transcriptional machinery
and/or RNA polymerase itself. These interactions
represent a key regulatory event in the multi-step
steroid receptor transactivation pathway.

IV. Nucleic Acid Encoding A SRC-1 Polypeptide.

15   Included within the scope of this invention
are the functional equivalents of the herein-described
isolated nucleic acid molecules. The degeneracy of the
genetic code permits substitution of certain codons by
other codons which specify the same amino acid and hence
20   would give rise to the same protein. The nucleic acid
sequence can vary substantially since, with the
exception of methionine and tryptophan, the known amino
acids can be coded for by more than one codon. Thus,
portions or all of the SRC-1 gene could be synthesized
25   to give a nucleic acid sequence significantly different
from that shown in Figure 1. The encoded amino acid
sequence thereof would, however, be preserved.

30

In addition, the nucleic acid sequence may comprise a nucleotide sequence which results from the addition, deletion or substitution of at least one nucleotide to the 5'-end and/or the 3'-end of the

5   nucleic acid formula shown in Figure 1 or a derivative thereof. Any nucleotide or polynucleotide may be used in this regard, provided that its addition, deletion or substitution does not alter the amino acid sequence of Figure 1 which is encoded by the nucleotide sequence.

10   For example, the present invention is intended to include any nucleic acid sequence resulting from the addition of ATG as an initiation codon at the 5'-end of the inventive nucleic acid sequence or its derivative, or from the addition of TTA, TAG or TGA as a termination

15   codon at the 3'-end of the inventive nucleotide sequence or its derivative. Moreover, the nucleic acid molecule of the present invention may, as necessary, have restriction endonuclease recognition sites added to its 5'-end and/or 3'-end.

20   Such functional alterations of a given nucleic acid sequence afford an opportunity to promote secretion and/or processing of heterologous proteins encoded by foreign nucleic acid sequences fused thereto. All variations of the nucleotide sequence of the SRC-1 genes

25   and fragments thereof permitted by the genetic code are, therefore, included in this invention.

      Further, it is possible to delete codons or to substitute one or more codons by codons other than degenerate codons to produce a structurally modified

31

polypeptide, but one which has substantially the same
utility or activity of the polypeptide produced by the
unmodified nucleic acid molecule. As recognized in the
art, the two polypeptides are functionally equivalent,

5   as are the two nucleic acid molecules which give rise to
their production, even though the differences between
the nucleic acid molecules are not related to degeneracy
of the genetic code.

V.   A Nucleic Acid Probe for the Detection of SRC-1.

10   A nucleic acid probe of the present invention
may be used to probe an appropriate chromosomal or cDNA
library by usual hybridization methods to obtain another
nucleic acid molecule of the present invention. A
chromosomal DNA or cDNA library may be prepared from

15   appropriate cells according to recognized methods in the
art (cf. "Molecular Cloning: A Laboratory Manual",
second edition, edited by Sambrook, Fritsch, & Maniatis,
Cold Spring Harbor Laboratory, 1989).

              In the alternative, chemical synthesis is

20   carried out in order to obtain nucleic acid probes
having nucleotide sequences which correspond to N-
terminal and C-terminal portions of the amino acid
sequence of the polypeptide of interest. Thus, the
synthesized nucleic acid probes may be used as primers

25   in a polymerase chain reaction (PCR) carried out in
accordance with recognized PCR techniques, essentially
according to PCR Protocols, "A Guide to Methods and
Applications", edited by Michael et al., Academic Press,

32

1990, utilizing the appropriate chromosomal or cDNA
library to obtain the fragment of the present invention.

One skilled in the art can readily design such
probes based on the sequence disclosed herein using
5   methods of computer alignment and sequence analysis
known in the art (cf. "Molecular Cloning: A Laboratory
Manual", second edition, edited by Sambrook, Fritsch, &
Maniatis, Cold Spring Harbor Laboratory, 1989). The
hybridization probes of the present invention can be
10   labeled by standard labeling techniques such as with a
radiolabel, enzyme label, fluorescent label, biotin-
avidin label, chemiluminescence, and the like. After
hybridization, the probes may be visualized using known
methods.

15   The nucleic acid probes of the present
invention include RNA, as well as DNA probes, such
probes being generated using techniques known in the
art. The nucleic acid probe may be immobilized on a
solid support. Examples of such solid supports include,
20   but are not limited to, plastics such as polycarbonate,
complex carbohydrates such as agarose and sepharose, and
acrylic resins, such as polyacrylamide and latex beads.
Techniques for coupling nucleic acid probes to such
solid supports are well known in the art.

25   The test samples suitable for nucleic acid
probing methods of the present invention include, for
example, cells or nucleic acid extracts of cells, or
biological fluids. The sample used in the above-
described methods will vary based on the assay format,

33

the detection method and the nature of the tissues,
cells or extracts to be assayed. Methods for preparing
nucleic acid extracts of cells are well known in the art
and can be readily adapted in order to obtain a sample
5   which is compatible with the method utilized.

VI.   **A Probe Based Method And Kit For**
       Detecting SRC-1.
              One method of detecting the presence of SRC-1
in a sample comprises a) contacting said sample with the
10   above-described nucleic acid probe, under conditions
such that hybridization occurs, and b) detecting the
presence of said probe bound to said nucleic acid
molecule. One skilled in the art would select the
nucleic acid probe according to techniques known in the
15   an as described above. Samples to be tested include but
should not be limited to RNA samples of human tissue.
              A kit for detecting the presence of SRC-1 in a
sample comprises at least one container means having
disposed therein the above-described nucleic acid probe.
20   The kit may further comprise other containers comprising
one or more of the following: wash reagents and reagents
capable of detecting the presence of bound nucleic acid
probe. Examples of detection reagents include, but are
not limited to radiolabelled probes, enzymatic labeled
25   probes (horseradish peroxidase, alkaline phosphatase),
and affinity labeled probes (biotin, avidin, or
steptavidin).

34

In detail, a compartmentalized kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or
5   paper. Such containers allow the efficient transfer of reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated and the agents or solutions of each container can be added in a quantitative fashion from
10   one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the probe or primers used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, and the
15   like), and containers which contain the reagents used to detect the hybridized probe, bound antibody, amplified product, or the like. One skilled in the art will readily recognize that the nucleic acid probes described in the present invention can readily be incorporated
20   into one of the established kit formats which are well known in the art.

VII. DNA Constructs Comprising a SRC-1
          Nucleic Acid Molecule and Cells
          Containing These Constructs.
25   The present invention also relates to a recombinant DNA molecule comprising, 5' to 3', a promoter effective to initiate transcription in a host cell and the above-described nucleic acid molecules. In

35

addition, the present invention relates to a recombinant
DNA molecule comprising a vector and an above-described
nucleic acid molecules. The present invention also
relates to a nucleic acid molecule comprising a
5   transcriptional region functional in a cell, a sequence
complimentary to an RNA sequence encoding an amino acid
sequence corresponding to the above-described
polypeptide, and a transcriptional termination region
functional in said cell. The above-described molecules
10   may be isolated and/or purified DNA molecules.

The present invention also relates to a cell
or organism that contains an above-described nucleic
acid molecule. The peptide may be purified from cells
which have been altered to express the peptide. A cell
15   is said to be "altered to express a desired peptide"
when the cell, through genetic manipulation, is made to
produce a protein which it normally does not produce or
which the cell normally produces at lower levels. One
skilled in the art can readily adapt procedures for
20   introducing and expressing either genomic, cDNA, or
synthetic sequences into either eukaryotic or
prokaryotic cells.

A nucleic acid molecule, such as DNA, is said
to be "capable of expressing" a polypeptide if it
25   contains nucleotide sequences which contain
transcriptional and translational regulatory information
and such sequences are "operably linked" to nucleotide
sequences which encode the polypeptide. An operable
linkage is a linkage in which the regulatory DNA

36

sequences and the DNA sequence sought to be expressed
are connected in such a way as to permit gene sequence
expression. The precise nature of the regulatory
regions needed for gene sequence expression may vary
5   from organism to organism, but shall in general include
a promoter region which, in prokaryotes, contains both
the promoter (which directs the initiation of RNA
transcription) as well as the DNA sequences which, when
transcribed into RNA, will signal synthesis initiation.
10   Such regions will normally include those 5'-non-coding
sequences involved with initiation of transcription and
translation, such as the TATA box, capping sequence,
CAAT sequence, and the like.

     If desired, the non-coding region 3' to the
15   sequence encoding an SRC-1 gene may be obtained by the
above-described methods. This region may be retained
for its transcriptional termination regulatory
sequences, such as termination and polyadenylation.
Thus, by retaining the 3'-region naturally contiguous to
20   the DNA sequence encoding an SRC-1 gene, the
transcriptional termination signals may be provided.
Where the transcriptional termination signals are not
satisfactorily functional in the expression host cell,
then a 3' region functional in the host cell may be
25   substituted.

     Two DNA sequences (such as a promoter region
sequence and an SRC-1 sequence) are said to be operably
linked if the nature of the linkage between the two DNA
sequences does not (1) result in the introduction of a

37

frame-shift mutation, (2) interfere with the ability of
the promoter region sequence to direct the transcription
of an SRC-1 gene sequence, or (3) interfere with the
ability of the an SRC-1 gene sequence to be transcribed

5   by the promoter region sequence. Thus, a promoter
region would be operably linked to a DNA sequence if the
promoter were capable of effecting transcription of that
DNA sequence. Thus, to express an SRC-1 gene,
transcriptional and translational signals recognized by

10   an appropriate host are necessary.

The present invention encompasses the
expression of the SRC-1 gene (or a functional derivative
thereof) in either prokaryotic or eukaryotic cells.
Prokaryotic hosts are, generally, very efficient and

15   convenient for the production of recombinant proteins
and are, therefore, one type of preferred expression
system for the SRC-1 gene. Prokaryotes most frequently
are represented by various strains of E. coli. However,
other microbial strains may also be used, including

20   other bacterial strains.

In prokaryotic systems, plasmid vectors that
contain replication sites and control sequences derived
from a species compatible with the host may be used.
Examples of suitable plasmid vectors may include pBR322,

25   pUC118, pUC119 and the like; suitable phage or
bacteriophage vectors may include γgt10, γgt11 and the
like; and suitable virus vectors may include pMAM-neo,
pKRC and the like. Preferably, the selected vector of

38

the present invention has the capacity to replicate in
the selected host cell.

Recognized prokaryotic hosts include bacteria
such as E. coil, Bacillus, Streptomyces, Pseudomonas,
5   Salmonella, Serratia, and the like. However, under such
conditions, the peptide will not be glycosylated. The
prokaryotic host must be compatible with the replicon
and control sequences in the expression plasmid.

To express SRC-1 (or a functional derivative
10   thereof) in a prokaryotic cell, it is necessary to
operably link the SRC-1 sequence to a functional
prokaryotic promoter. Such promoters may be either
constitutive or, more preferably, regulatable (i.e.,
inducible or derepressible). Examples of constitutive
15   promoters include the int promoter of bacteriophage λ,
the bla promoter of the β-lactamase gene sequence of
pBR322, and the CAT promoter of the chloramphenicol
acetyl transferase gene sequence of pPR325, and the
like. Examples of inducible prokaryotic promoters
20   include the major right and left promoters of
bacteriophage λ (P_L and P_R), the trp, recA, lacZ, lacI,
and gal promoters of E. coli, the α-amylase (Ulmanen et
at., J. Bacteriol. 162:176-182(1985)) and the ç-28-
specific promoters of B. subtilis (Gilman et at., Gene
25   sequence 32:11-20(1984)), the promoters of the
bacteriophages of Bacillus (Gryczan, In: The Molecular
Biology of the Bacilli, Academic Press, Inc., NY
(1982)), and Streptomyces promoters (Ward et at., Mol.
Gen. Genet. 203:468-478(1986)). Prokaryotic promoters

39

are reviewed by Glick (<u>J. Ind. Microbiol.</u> 1:277-
282(1987)); Cenatiempo (<u>Biochimie</u> 68:505-516(1986)); and
Gottesman (<u>Ann. Rev. Genet.</u> 18:415-442 (1984)).

Proper expression in a prokaryotic cell also

5   requires the presence of a ribosome binding site
upstream of the gene sequence-encoding sequence. Such
ribosome binding sites are disclosed, for example, by
Gold et at. (<u>Ann. Rev. Microbiol.</u> 35:365-404(1981)).
The selection of control sequences, expression vectors,

10   transformation methods, and the like, are dependent on
the type of host cell used to express the gene. As used
herein, "cell", "cell line", and "cell culture" may be
used interchangeably and all such designations include
progeny. Thus, the words "transformants" or

15   "transformed cells" include the primary subject cell and
cultures derived therefrom, without regard to the number
of transfers. It is also understood that all progeny
may not be precisely identical in DNA content, due to
deliberate or inadvertent mutations. However, as

20   defined, mutant progeny have the same functionality as
that of the originally transformed cell.

Host cells which may be used in the expression
systems of the present invention are not strictly
limited, provided that they are suitable for use in the

25   expression of the SRC-1 peptide of interest. Thus, any
primary human cell line such as those found in an ATTC
catalogue can be utilized. Suitable hosts may often
include eukaryotic cells. Preferred eukaryotic hosts
include, for example, yeast, fungi, insect cells,

40

mammalian cells either in vivo, or in tissue culture.
Mammalian cells which may be useful as hosts include
HeLa cells, cells of fibroblast origin such as VERO or
CHO-K1, or cells of lymphoid origin and their
5 derivatives. Preferred mammalian host cells include
SP2/0 and J558L, which may provide better capacities for
correct post-translational processing.

In addition, plant cells are also available as
hosts, and control sequences compatible with plant cells
10 are available, such as the cauliflower mosaic virus 35S
and 19S, and nopaline synthase promoter and
polyadenylation signal sequences. Another preferred
host is an insect cell, for example the Drosophila
larvae. Using insect cells as hosts, the Drosophila
15 alcohol dehydrogenase promoter can be used. Rubin,
Science 240:1453-1459(1988). Alternatively, baculovirus
vectors can be engineered to express large amounts of
SRC-1 in insects cells (Jasny, Science 238:1653 (1987);
Miller et al., In: Genetic Engineering (1986), Setlow,
20 J.K., et al., eds., Plenum, Vol. 8, pp. 277-297).

Any of a series of yeast gene sequence
expression systems can be utilized which incorporate
promoter and termination elements from the actively
expressed gene sequences coding for glycolytic enzymes
25 are produced in large quantities when yeast are grown in
mediums rich in glucose. Known glycolytic gene sequences
can also provide very efficient transcriptional control
signals. Yeast provides substantial advantages in that
it can also carry out post-translational peptide

41

modifications. A number of recombinant DNA strategies
exist which utilize strong promoter sequences and high
copy number of plasmids which can be utilized for
production of the desired proteins in yeast. Yeast

5   recognizes leader sequences on cloned mammalian gene
sequence products and secretes peptides bearing leader
sequences (i.e., pre-peptides). For a mammalian host,
several possible vector systems are available for the
expression of SRC-1.

10   A wide variety of transcriptional and
translational regulatory sequences may be employed,
depending upon the nature of the host. The
transcriptional and translational regulatory signals may
be derived from viral sources, such as adenovirus,

15   bovine papilloma virus, cytomegalovirus, papovavirus, or
the like, where the regulatory signals are associated
with a particular gene sequence which has a high level
of expression. Alternatively, promoters from mammalian
expression products, such as actin, collagen, myosin,

20   and the like, may be employed. Transcriptional
initiation regulatory signals may be selected which
allow for repression or activation, so that expression
of the gene sequences can be modulated. Of interest are
regulatory signals which are temperature-sensitive so

25   that by varying the temperature, expression can be
repressed or initiated, or are subject to chemical (such
as metabolite) regulation. Other regulatory signals
which may be utilized are described in U.S. Patent No.
5,364,791 and in U.S. Patent Application No. 07/939,246,

42

filed September 2, 1993, both of which are incorporated
herein by reference in their entirety including any
drawings.

Expression of SRC-1 in eukaryotic hosts
5   requires the use of eukaryotic regulatory regions. Such
regions will, in general, include a promoter region
sufficient to direct the initiation of RNA synthesis.
Preferred eukaryotic promoters include, for example, the
promoter of the mouse metallothionein I gene sequence
10   (Hamer et al., J. Mol. Appl. Gen. 1:273-288(1982)); the
TK promoter of Herpes virus (McKnight, Cell 31:355-365
(1982)); the SV40 early promoter (Benoist et al., Nature
(London) 290:304-310(1981)); the yeast gal4 gene
sequence promoter (Johnston et al., Proc. Natl. Acad.
15   Sci. (USA) 79:6971-6975(1982); Silver et al., Proc.
Natl. Acad. Sci. (USA) 81:5951-5955 (1984)).

Translation of eukaryotic mRNA is initiated at
the codon which encodes the first methionine. For this
reason, it is preferable to ensure that the linkage
20   between a eukaryotic promoter and a DNA sequence which
encodes SRC-1 (or a functional derivative thereof) does
not contain any intervening codons which are capable of
encoding a methionine (i.e., AUG). The presence of such
codons results either in a formation of a fusion protein
25   (if the AUG codon is in the same reading frame as the
SRC-1 coding sequence) or a frame-shift mutation (if the
AUG codon is not in the same reading frame as the SRC-1
coding sequence).

43

A SRC-1 nucleic acid molecule and an operably
linked promoter may be introduced into a recipient
prokaryotic or eukaryotic cell either as a
nonreplicating DNA (or RNA) molecule, which may either

5   be a linear molecule or, more preferably, a closed
covalent circular molecule. Since such molecules are
incapable of autonomous replication, the expression of
the gene may occur through the transient expression of
the introduced sequence. Alternatively, permanent

10   expression may occur through the integration of the
introduced DNA sequence into the host chromosome.

A vector may be employed which is capable of
integrating the desired gene sequences into the host
cell chromosome. Cells which have stably integrated the

15   introduced DNA into their chromosomes can be selected by
also introducing one or more markers which allow for
selection of host cells which contain the expression
vector. The marker may provide for prototrophy to an
auxotrophic host, biocide resistance, e.g., antibiotics,

20   or heavy metals, such as copper, or the like. The
selectable marker gene sequence can either be directly
linked to the DNA gene sequences to be expressed, or
introduced into the same cell by co-transfection.
Additional elements may also be needed for optimal

25   synthesis of single chain binding protein mRNA. These
elements may include splice signals, as well as
transcription promoters, enhancers, and termination
signals. cDNA expression vectors incorporating such

44

elements include those described by Okayama, _Molec.
Cell. Biol._ *3*:280(1983).

The introduced nucleic acid molecule can be
incorporated into a plasmid or viral vector capable of
5   autonomous replication in the recipient host. Any of a
wide variety of vectors may be employed for this
purpose. Factors of importance in selecting a
particular plasmid or viral vector include: the ease
with which recipient cells that contain the vector may
10   be recognized and selected from those recipient cells
which do not contain the vector; the number of copies of
the vector which are desired in a particular host; and
whether it is desirable to be able to "shuttle" the
vector between host cells of different species.

15   Preferred prokaryotic vectors include plasmids
such as those capable of replication in E. coil (such
as, for example, pBR322, ColEl, pSC101, pACYC 184, nVX.
Such plasmids are, for example, disclosed by Sambrook
(cf. "Molecular Cloning: A Laboratory Manual", second
20   edition, edited by Sambrook, Fritsch, & Maniatis, Cold
Spring Harbor Laboratory, (1989)). Bacillus plasmids
include pC194, pC221, pT127, and the like. Such plasmids
are disclosed by Gryczan (In: The Molecular Biology of
the Bacilli, Academic Press, NY (1982), pp. 307-329).
25   Suitable Streptomyces plasmids include pIJ101 (Kendall
et al., _J. Bacteriol._ *169*:4177-4183 (1987)), and
streptomyces bacteriophages such as φC31 (Chater et al.,
In: Sixth International Symposium on Actinomycetales
Biology, Akademiai Kaido, Budapest, Hungary (1986), pp.

45

45-54). Pseudomonas plasmids are reviewed by John et al.
(Rev. Infect. Dis. 8:693-704(1986)), and Izaki (Jpn. J.
Bacteriol. 33:729-742(1978)).

Preferred eukaryotic plasmids include, for
5   example, BPV, vaccinia, SV40, 2-micron circle,
adenovirus, retrovirus, and the like, or their
derivatives. Such plasmids are well known in the art
(Botstein et al., Miami Wntr. Symp. 19:265-274(1982);
Broach, In: "The Molecular Biology of the Yeast
10   Saccharomyces: Life Cycle and Inheritance", Cold Spring
Harbor Laboratory, Cold Spring Harbor, NY, p. 445-470
(1981); Broach, Cell 28:203-204 (1982); Bollon et at.,
J. Ctin. Hematol. Oncol. 10:39-48 (1980); Maniatis, In:
Cell Biology: A Comprehensive Treatise, Vol. 3, Gene
15   Sequence Expression, Academic Press, NY, pp. 563-
608(1980).

Once the vector or nucleic acid molecule
containing the construct(s) has been prepared for
expression, the DNA construct(s) may be introduced into
20   an appropriate host cell by any of a variety of suitable
means, i.e., transformation, transfection, conjugation,
protoplast fusion, electroporation, particle gun
technology, calcium phosphate-precipitation, direct
microinjection, and the like. After the introduction of
25   the vector, recipient cells are grown in a selective
medium, which selects for the growth of vector-
containing cells. Expression of the cloned gene
molecule(s) results in the production of SRC-1 or
fragments thereof. This can take place in the

46

transformed cells as such, or following the induction of
these cells to differentiate (for example, by
administration of bromodeoxyuracil to neuroblastoma
cells or the like). A variety of incubation conditions

5   can be used to form the peptide of the present
invention. The most preferred conditions are those which
mimic physiological conditions.

VIII.   **Purified SRC-1 Polypeptides**

In another aspect the invention features an

10   isolated, enriched, or purified SRC-1 polypeptide.

By "isolated" in reference to a polypeptide is
meant a polymer of 2 (preferably 7, more preferably 13,
most prefereably 25) or more amino acids conjugated to
each other, including polypeptides that are isolated

15   from a natural source or that are synthesized. In
certain aspects longer polypeptides are preferred, such
as those with 402, 407, 413, or 425 contiguous amino
acids set forth in Figure 1. The isolated polypeptides
of the present invention are unique in the sense that

20   they are not found in a pure or separated state in
nature. Use of the term "isolated" indicates that a
naturally occurring sequence has been removed from its
normal cellular environment. Thus, the sequence may be
in a cell-free solution or placed in a different

25   cellular environment. The term does not imply that the
sequence is the only amino acid chain present, but that
it is essentially free (about 90 - 95% pure at least) of
non-amino acid material naturally associated with it.

47

By the use of the term "enriched" in reference
to a polypeptide is meant that the specific amino acid
sequence constitutes a significantly higher fraction (2
- 5 fold) of the total of amino acids present in the

5   cells or solution of interest than in normal or diseased
cells or in the cells from which the sequence was taken.
This could be caused by a person by preferential
reduction in the amount of other amino acids present, or
by a preferential increase in the amount of the specific

10   amino acid sequence of interest, or by a combination of
the two. However, it should be noted that enriched does
not imply that there are no other amino acid sequences
present, just that the relative amount of the sequence
of interest has been significantly increased. The term

15   significant here is used to indicate that the level of
increase is useful to the person making such an
increase, and generally means an increase relative to
other amino acids of about at least 2 fold, more
preferably at least 5 to 10 fold or even more. The term

20   also does not imply that there is no amino acid from
other sources. The other source amino acid may, for
example, comprise amino acid encoded by a yeast or
bacterial genome, or a cloning vector such as pUC19.
The term is meant to cover only those situations in

25   which man has intervened to elevate the proportion of
the desired nucleic acid.

It is also advantageous for some purposes that
an amino acid sequence be in purified form. The term
"purified" in reference to a polypeptide does not

48

require absolute purity (such as a homogeneous
preparation); instead, it represents an indication that
the sequence is relatively purer than in the natural
environment (compared to the natural level this level

5   should be at least 2-5 fold greater, e.g., in terms of
mg/ml). Purification of at least one order of
magnitude, preferably two or three orders, and more
preferably four or five orders of magnitude is expressly
contemplated. The substance is preferably free of

10   contamination at a functionally significant level, for
example 90%, 95%, or 99% pure.

In preferred embodiments the SRC-1 polypeptide
contains at least 25, 30, 35, 40, 50, 425, 430, 435,
440, or 450 contiguous amino acids of the full-length

15   sequence set forth in Figure 1, or a functional
derivative thereof.

In another aspect, the invention describes a
polypeptide comprising a recombinant SRC-1 polypeptide
or a unique fragment thereof. By "unique fragment," is

20   meant an amino acid sequence present in a full-length
SRC-1 polypeptide that is not present in any other
naturally occurring polypeptide. Preferably, such a
sequence comprises 6 contiguous amino acids present in
the full sequence. More preferably, such a sequence

25   comprises 12 contiguous amino acids present in the full
sequence. Even more preferably, such a sequence
comprises 18 contiguous amino acids present in the full
sequence.

49

By "recombinant SRC-1 polypeptide" is meant to include a polypeptide produced by recombinant DNA techniques such that it is distinct from a naturally occurring polypeptide either in its location (e.g., 5 present in a different cell or tissue than found in nature), purity or structure. Generally, such a recombinant polypeptide will be present in a cell in an amount different from that normally observed in nature.

10 A variety of methodologies known in the art can be utilized to obtain the peptide of the present invention. The peptide may be purified from tissues or cells which naturally produce the peptide. Alternatively, the above-described isolated nucleic acid 15 fragments could be used to expressed the SRC-1 protein in any organism. The samples of the present invention include cells, protein extracts or membrane extracts of cells, or biological fluids. The sample will vary based on the assay format, the detection method and the nature 20 of the tissues, cells or extracts used as the sample.

Any eukaryotic organism can be used as a source for the peptide of the invention, as long as the source organism naturally contains such a peptide. As used herein, "source organism" refers to the original 25 organism from which the amino acid sequence of the subunit is derived, regardless of the organism the subunit is expressed in and ultimately isolated from.

One skilled in the art can readily follow known methods for isolating proteins in order to obtain

50

the peptide free of natural contaminants. These
include, but are not limited to: size-exclusion
chromatography, HPLC, ion-exchange chromatography, and
immuno-affinity chromatography.

5   IX. An Antibody Having Binding Affinity
         To A SRC-1 Polypeptide And A
         Hybridoma Containing the Antibody.

              In yet another aspect the invention features
     an antibody (e.g., a monoclonal or polyclonal antibody)
10   having specific binding affinity to a SRC-1 polypeptide.
     The antibody contains a sequence of amino acids that is
     able to specifically bind to a SRC-1 polypeptide. By
     "specific binding affinity" is meant that the antibody
     binds to SRC-1 polypeptides with greater affinity than
15   it binds to other polypeptides under specified
     conditions.

              Antibodies having specific binding affinity to
     a SRC-1 polypeptide may be used in methods for detecting
     the presence and/or amount of a SRC-1 polypeptide in a
20   sample by contacting the sample with the antibody under
     conditions such that an immunocomplex forms and
     detecting the presence and/or amount of the antibody
     conjugated to the SRC-1 polypeptide. Diagnostic kits
     for performing such methods may be constructed to
25   include a first container means containing the antibody
     and a second container means having a conjugate of a
     binding partner of the antibody and a label.

51

In another aspect the invention features a hybridoma which produces an antibody having specific binding affinity to a SRC-1 polypeptide. By "hybridoma" is meant an immortalized cell line which is capable of

5   secreting an antibody, for example a SRC-1 antibody. In preferred embodiments the SRC-1 antibody comprises a sequence of amino acids that is able to specifically bind a SRC-1 polypeptide.

The present invention relates to an antibody

10   having binding affinity to a SRC-1 polypeptide. The polypeptide may have the amino acid sequence set forth in Figure 1, or functional derivative thereof, or at least 9 contiguous amino acids thereof (preferably, at least 20, 30, 35, or 40 contiguous amino acids thereof).

15   The present invention also relates to an antibody having specific binding affinity to an SRC-1 polypeptide. Such an antibody may be isolated by comparing its binding affinity to a SRC-1 polypeptide with its binding affinity to another polypeptide. Those

20   which bind selectively to SRC-1 would be chosen for use in methods requiring a distinction between SRC-1 and other polypeptides. Such methods could include, but should not be limited to, the analysis of altered SRC-1 expression in tissue containing other polypeptides.

25   The SRC-1 proteins of the present invention can be used in a variety of procedures and methods, such as for the generation of antibodies, for use in identifying pharmaceutical compositions, and for studying DNA/protein interaction.

52

The SRC-1 peptide of the present invention can be used to produce antibodies or hybridomas. One skilled in the art will recognize that if an antibody is desired, such a peptide would be generated as described herein and used as an immunogen. The antibodies of the present invention include monoclonal and polyclonal antibodies, as well fragments of these antibodies, and humanized forms. Humanized forms of the antibodies of the present invention may be generated using one of the procedures known in the art such as chimerization or CDR grafting. The present invention also relates to a hybridoma which produces the above-described monoclonal antibody, or binding fragment thereof. A hybridoma is an immortalized cell line which is capable of secreting a specific monoclonal antibody.

In general, techniques for preparing monoclonal antibodies and hybridomas are well known in the art (Campbell, "Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology," Elsevier Science Publishers, Amsterdam, The Netherlands (1984); St. Groth et al., J. Immunol. Methods 35:1-21(1980)). Any animal (mouse, rabbit, and the like) which is known to produce antibodies can be immunized with the selected polypeptide. Methods for immunization are well known in the art. Such methods include subcutaneous or intraperitoneal injection of the polypeptide. One skilled in the art will recognize that the amount of polypeptide used for immunization will vary based on the animal which is immunized, the

53

antigenicity of the polypeptide and the site of
injection.

The polypeptide may be modified or
administered in an adjuvant in order to increase the
5   peptide antigenicity. Methods of increasing the
antigenicity of a polypeptide are well known in the art.
Such procedures include coupling the antigen with a
heterologous protein (such as globulin or β-
galactosidase) or through the inclusion of an adjuvant
10   during immunization.

For monoclonal antibodies, spleen cells from
the immunized animals are removed, fused with myeloma
cells, such as SP2/0-Ag14 myeloma cells, and allowed to
become monoclonal antibody producing hybridoma cells.
15   Any one of a number of methods well known in the art can
be used to identify the hybridoma cell which produces an
antibody with the desired characteristics. These include
screening the hybridomas with an ELISA assay, western
blot analysis, or radioimmunoassay (Lutz et al., Exp.
20   Cell Res. 175:109-124(1988)). Hybridomas secreting the
desired antibodies are cloned and the class and subclass
is determined using procedures known in the art
(Campbell, "Monoclonal Antibody Technology: Laboratory
Techniques in Biochemistry and Molecular Biology", supra
25   (1984)).

For polyclonal antibodies, antibody containing
antisera is isolated from the immunized animal and is
screened for the presence of antibodies with the desired
specificity using one of the above-described procedures.

54

The above-described antibodies may be detectably
labeled. Antibodies can be detectably labeled through
the use of radioisotopes, affinity labels (such as
biotin, avidin, and the like), enzymatic labels (such as
5   horse radish peroxidase, alkaline phosphatase, and the
like) fluorescent labels (such as FITC or rhodamine, and
the like), paramagnetic atoms, and the like. Procedures
for accomplishing such labeling are well-known in the
art, for example, see (Stemberger et al., _J. Histochem._
10   _Cytochem._ 18:315 (1970); Bayer et at., _Meth. Enzym._
62:308 (1979); Engval et al., _Immunot._ 109:129(1972);
Goding, _J. Immunol. Meth._ 13:215(1976)). The labeled
antibodies of the present invention can be used for in
vitro, in vivo, and in situ assays to identify cells or
15   tissues which express a specific peptide.

The above-described antibodies may also be
immobilized on a solid support. Examples of such solid
supports include plastics such as polycarbonate, complex
carbohydrates such as agarose and sepharose, acrylic
20   resins and such as polyacrylamide and latex beads.
Techniques for coupling antibodies to such solid
supports are well known in the art (Weir et al.,
"Handbook of Experimental Immunology" 4th Ed., Blackwell
Scientific Publications, Oxford, England, Chapter
25   10(1986); Jacoby et al., _Meth. Enzym._ 34, Academic
Press, N.Y. (1974)). The immobilized antibodies of the
present invention can be used for in vitro, in vivo, and
in situ assays as well as in immunochromotography.

55

Furthermore, one skilled in the art can
readily adapt currently available procedures, as well as
the techniques, methods and kits disclosed above with
regard to antibodies, to generate peptides capable of
5   binding to a specific peptide sequence in order to
generate rationally designed antipeptide peptides, for
example see Hurby et al., "Application of Synthetic
Peptides: Antisense Peptides", In Synthetic Peptides, A
User's Guide, W.H. Freeman, NY, pp. 289-307(1992), and
10   Kaspczak et al., Biochemistry 28:9230-8(1989).

Anti-peptide peptides can be generated by
replacing the basic amino acid residues found in the
SRC-1 peptide sequence with acidic residues, while
maintaining hydrophobic and uncharged polar groups. For
15   example, lysine, arginine, and/or histidine residues are
replaced with aspartic acid or glutamic acid and
glutamic acid residues are replaced by lysine, arginine
or histidine.

X.   An Antibody Based Method And Kit
20   **For Detecting SRC-1.**

The present invention encompasses a method of
detecting an SRC-1 polypeptide in a sample, comprising:
a) contacting the sample with an above-described
antibody, under conditions such that immunocomplexes
25   form, and b) detecting the presence of said antibody
bound to the polypeptide. In detail, the methods
comprise incubating a test sample with one or more of
the antibodies of the present invention and assaying

56

whether the antibody binds to the test sample. Altered levels of SRC-1 in a sample as compared to normal levels may indicate disease.

Conditions for incubating an antibody with a
5   test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the antibody used in the assay. One skilled in the art will recognize that any one of the commonly available immunological assay
10   formats (such as radioimmunoassays, enzyme-linked immunosorbent assays, diffusion based Ouchterlony, or rocket immunofluorescent assays) can readily be adapted to employ the antibodies of the present invention. Examples of such assays can be found in Chard, "An
15   Introduction to Radioimmunoassay and Related Techniques" Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock et al., "Techniques in Immunocytochemistry," Academic Press, Orlando, FL Vol. 1(1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen,
20   "Practice and Theory of Enzyme Immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology," Elsevier Science Publishers, Amsterdam, The Netherlands (1985).

The immunological assay test samples of the
25   present invention include cells, protein or membrane extracts of cells, or biological fluids such as blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues,

57

cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is

5   capable with the system utilized.

A kit contains all the necessary reagents to carry out the previously described methods of detection. The kit may comprise: i) a first container means containing an above-described antibody, and ii) second

10   container means containing a conjugate comprising a binding partner of the antibody and a label. In another preferred embodiment, the kit further comprises one or more other containers comprising one or more of the following: wash reagents and reagents capable of

15   detecting the presence of bound antibodies.

Examples of detection reagents include, but are not limited to, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the chromophoric, enzymatic, or antibody binding reagents

20   which are capable of reacting with the labeled antibody. The compartmentalized kit may be as described above for nucleic acid probe kits. One skilled in the art will readily recognize that the antibodies described in the present invention can readily be incorporated into one

25   of the established kit formats which are well known in the art.

58

XI.   Isolation of Compounds Which
Interact With SRC-1.

The present invention also relates to a method
of detecting a compound capable of binding to a SRC-1
5   polypeptide comprising incubating the compound with SRC-
1 and detecting the presence of the compound bound to
SRC-1. The compound may be present within a complex
mixture, for example, serum, body fluid, or cell
extracts.

10   The present invention also relates to a method
of detecting an agonist or antagonist of SRC-1 activity
comprising incubating cells that produce SRC-1 in the
presence of a compound and detecting changes in the
level of SRC-1 activity. The compounds thus identified
15   would produce a change in activity indicative of the
presence of the compound. The compound may be present
within a complex mixture, for example, serum, body
fluid, or cell extracts. Once the compound is
identified it can be isolated using techniques well
20   known in the art.

The present invention also encompasses a
method of agonizing (stimulating) or antagonizing SRC-1
associated activity in a mammal comprising administering
to said mammal an agonist or antagonist to SRC-1 in an
25   amount sufficient to effect said agonism or antagonism.
A method of treating diseases in a mammal with an
agonist or antagonist of SRC-1 activity comprising
administering the agonist or antagonist to a mammal in
an amount sufficient to agonize or antagonize SRC-1

59

associated functions is also encompassed in the present application.

XII. Transgenic Animals.

A variety of methods are available for the
5   production of transgenic animals associated with this invention. DNA can be injected into the pronucleus of a fertilized egg before fusion of the male and female pronuclei, or injected into the nucleus of an embryonic cell (e.g., the nucleus of a two-cell embryo) following
10   the initiation of cell division (Brinster et al., Proc. Nat. Acad. Sci. USA 82: 4438-4442 (1985)). Embryos can be infected with viruses, especially retroviruses, modified to carry inorganic-ion receptor nucleotide sequences of the invention.
15   Pluripotent stem cells derived from the inner cell mass of the embryo and stabilized in culture can be manipulated in culture to incorporate nucleotide sequences of the invention. A transgenic animal can be produced from such cells through implantation into a
20   blastocyst that is implanted into a foster mother and allowed to come to term. Animals suitable for transgenic experiments can be obtained from standard commercial sources such as Charles River (Wilmington, MA), Taconic (Germantown, NY), Harlan Sprague Dawley
25   (Indianapolis, IN), etc.

The procedures for manipulation of the rodent embryo and for microinjection of DNA into the pronucleus of the zygote are well known to those of ordinary skill

60

in the art (Hogan et al., supra). Microinjection
procedures for fish, amphibian eggs and birds are
detailed in Houdebine and Chourrout, Experientia 47:
897-905 (1991). Other procedures for introduction of

5   DNA into tissues of animals are described in U.S. Patent
No., 4,945,050 (Sandford et al., July 30, 1990).

By way of example only, to prepare a
transgenic mouse, female mice are induced to
superovulate. Females are placed with males, and the

10   mated females are sacrificed by CO_2 asphyxiation or
cervical dislocation and embryos are recovered from
excised oviducts. Surrounding cumulus cells are
removed. Pronuclear embryos are then washed and stored
until the time of injection. Randomly cycling adult

15   female mice are paired with vasectomized males.
Recipient females are mated at the same time as donor
females. Embryos then are transferred surgically. The
procedure for generating transgenic rats is similar to
that of mice. See Hammer et al., Cell 63:1099-1112

20   (1990).

Methods for the culturing of embryonic stem
(ES) cells and the subsequent production of transgenic
animals by the introduction of DNA into ES cells using
methods such as electroporation, calcium phosphate/DNA

25   precipitation and direct injection also are well known
to those of ordinary skill in the art. See, for
example, Teratocarcinomas and Embryonic Stem Cells. A
Practical Approach, E.J. Robertson, ed., IRL Press
(1987).

61

In cases involving random gene integration, a clone containing the sequence(s) of the invention is co-transfected with a gene encoding resistance. Alternatively, the gene encoding neomycin resistance is

5   physically linked to the sequence(s) of the invention. Transfection and isolation of desired clones are carried out by any one of several methods well known to those of ordinary skill in the art (E.J. Robertson, supra).

DNA molecules introduced into ES cells can

10   also be integrated into the chromosome through the process of homologous recombination. Capecchi, Science 244: 1288-1292 (1989). Methods for positive selection of the recombination event (i.e., neo resistance) and dual positive-negative selection (i.e., neo resistance

15   and gancyclovir resistance) and the subsequent identification of the desired clones by PCR have been described by Capecchi, supra and Joyner et al., Nature 338: 153-156 (1989), the teachings of which are incorporated herein. The final phase of the procedure

20   is to inject targeted ES cells into blastocysts and to transfer the blastocysts into pseudopregnant females. The resulting chimeric animals are bred and the offspring are analyzed by Southern blotting to identify individuals that carry the transgene. Procedures for

25   the production of non-rodent mammals and other animals have been discussed by others. See Houdebine and Chourrout, supra; Pursel et al., Science 244:1281-1288 (1989); and Simms et al., Bio/Technology 6:179-183 (1988).

62

XIII.   Gene Therapy

SRC-1 or its genetic sequences will also be useful in gene therapy (reviewed in Miller, Nature 357:455-460, (1992). Miller states that advances have
5   resulted in practical approaches to human gene therapy that have demonstrated positive initial results. The basic science of gene therapy is described in Mulligan, Science 260:926-931, (1993).

In one preferred embodiment, an expression
10   vector containing the SRC-1 coding sequence is inserted into cells, the cells are grown in vitro and then infused in large numbers into patients. In another preferred embodiment, a DNA segment containing a promoter of choice (for example a strong promoter) is
15   transferred into cells containing an endogenous SRC-1 in such a manner that the promoter segment enhances expression of the endogenous SRC-1 gene (for example, the promoter segment is transferred to the cell such that it becomes directly linked to the endogenous SRC-1
20   gene).

The gene therapy may involve the use of an adenovirus containing SRC-1 cDNA targeted to a tumor, systemic SRC-1 increase by implantation of engineered cells, injection with SRC-1 virus, or injection of naked
25   SRC-1 DNA into appropriate tissues.

Target cell populations may be modified by introducing altered forms of one or more components of the protein complexes in order to modulate the activity of such complexes. For example, by reducing or inhibit-

63

ing a complex component activity within target cells, an

abnormal signal transduction event(s) leading to a

condition may be decreased, inhibited, or reversed.

Deletion or missense mutants of a component, that retain

5   the ability to interact with other components of the

protein complexes but cannot function in signal trans-

duction may be used to inhibit an abnormal, deleterious

signal transduction event.

Expression vectors derived from viruses such

10   as retroviruses, vaccinia virus, adenovirus, adeno-

associated virus, herpes viruses, several RNA viruses,

or bovine papilloma virus, may be used for delivery of

nucleotide sequences (e.g., cDNA) encoding recombinant

SRC-1 protein into the targeted cell population (e.g.,

15   tumor cells). Methods which are well known to those

skilled in the art can be used to construct recombinant

viral vectors containing coding sequences. See, for

example, the techniques described in Maniatis et al.,

Molecular Cloning: A Laboratory Manual, Cold Spring

20   Harbor Laboratory, N.Y. (1989), and in Ausubel et al.,

Current Protocols in Molecular Biology, Greene

Publishing Associates and Wiley Interscience, N.Y.

(1989). Alternatively, recombinant nucleic acid mole-

cules encoding protein sequences can be used as naked

25   DNA or in reconstituted system e.g., liposomes or other

lipid systems for delivery to target cells (See e.g.,

Felgner et al., Nature 337:387-8, 1989). Several other

methods for the direct transfer of plasmid DNA into

cells exist for use in human gene therapy and involve

64

targeting the DNA to receptors on cells by complexing
the plasmid DNA to proteins. See, Miller, supra.

In its simplest form, gene transfer can be
performed by simply injecting minute amounts of DNA into

5   the nucleus of a cell, through a process of
microinjection. Capecchi MR, Cell 22:479-88 (1980).
Once recombinant genes are introduced into a cell, they
can be recognized by the cells normal mechanisms for
transcription and translation, and a gene product will

10   be expressed. Other methods have also been attempted
for introducing DNA into larger numbers of cells. These
methods include: transfection, wherein DNA is
precipitated with $CaPO_4$ and taken into cells by
pinocytosis (Chen C. and Okayama H, Mol. Cell Biol.

15   7:2745-52 (1987)); electroporation, wherein cells are
exposed to large voltage pulses to introduce holes into
the membrane (Chu G. et al., Nucleic Acids Res.,
15:1311-26 (1987)); lipofection/liposome fusion, wherein
DNA is packaged into lipophilic vesicles which fuse with

20   a target cell (Felgner PL., et al., Proc. Natl. Acad.
Sci. USA. 84:7413-7 (1987)); and particle bombardment
using DNA bound to small projectiles (Yang NS. et al.,
Proc. Natl. Acad. Sci. 87:9568-72 (1990)). Another
method for introducing DNA into cells is to couple the

25   DNA to chemically modified proteins.

It has also been shown that adenovirus
proteins are capable of destabilizing endosomes and
enhancing the uptake of DNA into cells. The admixture
of adenovirus to solutions containing DNA complexes, or

65

the binding of DNA to polylysine covalently attached to
adenovirus using protein crosslinking agents
substantially improves the uptake and expression of the
recombinant gene. Curiel DT et al., <u>Am. J. Respir.</u>

5   <u>Cell. Mol. Biol.</u>, 6:247-52 (1992).

As used herein "gene transfer" means the
process of introducing a foreign nucleic acid molecule
into a cell. Gene transfer is commonly performed to
enable the expression of a particular product encoded by

10   the gene. The product may include a protein,
polypeptide, anti-sense DNA or RNA, or enzymatically
active RNA. Gene transfer can be performed in cultured
cells or by direct administration into animals.
Generally gene transfer involves the process of nucleic

15   acid contact with a target cell by non-specific or
receptor mediated interactions, uptake of nucleic acid
into the cell through the membrane or by endocytosis,
and release of nucleic acid into the cytoplasm from the
plasma membrane or endosome. Expression may require, in

20   addition, movement of the nucleic acid into the nucleus
of the cell and binding to appropriate nuclear factors
for transcription.

As used herein "gene therapy" is a form of
gene transfer and is included within the definition of

25   gene transfer as used herein and specifically refers to
gene transfer to express a therapeutic product from a
cell *in vivo* or *in vitro*. Gene transfer can be
performed *ex vivo* on cells which are then transplanted
into a patient, or can be performed by direct

administration of the nucleic acid or nucleic acid-
protein complex into the patient.

In another preferred embodiment, a vector
having nucleic acid sequences encoding SRC-1 is provided
5   in which the nucleic acid sequence is expressed only in
specific tissue. Methods of achieving tissue-specific
gene expression as set forth in International
Publication No. WO 93/09236, filed November 3, 1992 and
published May 13, 1993.

10   In all of the preceding vectors set forth
above, a further aspect of the invention is that the
nucleic acid sequence contained in the vector may
include additions, deletions or modifications to some or
all of the sequence of the nucleic acid, as defined
15   above.

In another preferred embodiment, a method of
gene replacement is set forth. "Gene replacement" as
used herein means supplying a nucleic acid sequence
which is capable of being expressed *in vivo* in an animal
20   and thereby providing or augmenting the function of an
endogenous gene which is missing or defective in the
animal.

The nucleic acid sequence encoding SRC-1 can
be administered prophylactically, or to patients having
25   a disorder listed above, *e.g.*, by exogenous delivery of
the nucleic acid sequence encoding SRC-1 as naked DNA,
DNA associated with specific carriers, or in a nucleic
acid expression vector to a desired tissue by means of
an appropriate delivery vehicle, *e.g.*, a liposome, by

67

use of iontophoresis, electroporation and other
pharmacologically approved methods of delivery. Routes
of administration may include intramuscular,
intravenous, aerosol, oral (tablet or pill form),

5   topical, systemic, ocular, as a suppository,
intraperitoneal and/or intrathecal.

Some methods of delivery that may be used
include:

   a.   encapsulation in liposomes,

10   b.   transduction by retroviral vectors,

   c.   localization to nuclear compartment
         utilizing nuclear targeting site found on
         most nuclear proteins,

   d.   transfection of cells *ex vivo* with subse-

15   quent reimplantation or administration of
         the transfected cells,

   e.   a DNA transporter system.

A SRC-1 nucleic acid sequence may be
administered utilizing an *ex vivo* approach whereby cells

20   are removed from an animal, transduced with the SRC-1
nucleic acid sequence and reimplanted into the animal.
The liver can be accessed by an *ex vivo* approach by
removing hepatocytes from an animal, transducing the
hepatocytes *in vitro* with the SRC-1 nucleic acid

25   sequence and reimplanting them into the animal (*e.g.*, as
described for rabbits by Chowdhury et al, Science 254:
1802-1805, 1991, or in humans by Wilson, Hum. Gene Ther.
3: 179-222, 1992) incorporated herein by reference.

68

Many nonviral techniques for the delivery of a
SRC-1 nucleic acid sequence into a cell can be used,
including direct naked DNA uptake (*e.g.*, Wolff et al.,
Science 247: 1465-1468, 1990), receptor-mediated DNA

5   uptake, *e.g.*, using DNA coupled to asialoorosomucoid
which is taken up by the asialoglycoprotein receptor in
the liver (Wu and Wu, J. Biol. Chem. 262: 4429-4432,
1987; Wu et al., J. Biol. Chem. 266: 14338-14342, 1991),
and liposome-mediated delivery (*e.g.*, Kaneda et al.,

10   Expt. Cell Res. 173: 56-69, 1987; Kaneda et al., Science
243: 375-378, 1989; Zhu et al., Science 261: 209-211,
1993). Many of these physical methods can be combined
with one another and with viral techniques; enhancement
of receptor-mediated DNA uptake can be effected, for

15   example, by combining its use with adenovirus (Curiel et
al., Proc. Natl. Acad. Sci. USA 88: 8850-8854, 1991;
Cristiano et al., Proc. Natl. Acad. Sci. USA 90: 2122-
2126, 1993).

The SRC-1 or nucleic acid encoding a SRC-1

20   polypeptide or protein may also be administered via an
implanted device that provides a support for growing
cells. Thus, the cells may remain in the implanted
device and still provide the useful and therapeutic
agents of the present invention.

25   The term "protein" refers to a compound formed
of 5-50 or more amino acids joined together by peptide
bonds. An "amino acid" is a subunit that is polymerized
to form proteins and there are twenty amino acids that
are universally found in proteins. The general formula

69

for an amino acid is H_2N-CHR-COOH, in which the R group
can be anything from a hydrogen atom (as in the amino
acid glycine) to a complex ring (as in the amino acid
tryptophan).

5   **XV. Derivatives**

Also provided herein are functional deriva-
tives of SRC-1. By "functional derivative" is meant a
"chemical derivative," "fragment," "variant," "chimera,"
or "hybrid", which terms are defined below. A

10   functional derivative retains at least a portion of the
function of the protein, for example reactivity with an
antibody specific for the complex, enzymatic activity or
binding activity mediated through noncatalytic domains,
which permits its utility in accordance with the present

15   invention.

A "chemical derivative" contains additional
chemical moieties not normally a part of the protein.
Covalent modifications of the protein or peptide are
included within the scope of this invention. Such

20   modifications may be introduced into the molecule by
reacting targeted amino acid residues of the peptide
with an organic derivatizing agent that is capable of
reacting with selected side chains or terminal residues,
as described below.

25   Cysteinyl residues most commonly are reacted
with alpha-haloacetates (and corresponding amines), such
as chloroacetic acid or chloroacetamide, to give car-
boxymethyl or carboxyamidomethyl derivatives. Cysteinyl

residues also are derivatized by reaction with bromotri-
fluoroacetone, chloroacetyl phosphate, N-alkylmalei-
mides, 3-nitro-2-pyridyl disulfide, methyl 2-pyridyl
disulfide, p-chloromercuribenzoate, 2-chloromercuri-4-
5   nitrophenol, or chloro-7-nitrobenzo-2-oxa-1,3-diazole.

Histidyl residues are derivatized by reaction
with diethylprocarbonate at pH 5.5-7.0 because this
agent is relatively specific for the histidyl side
chain. Para-bromophenacyl bromide also is useful; the
10   reaction is preferably performed in 0.1 M sodium
cacodylate at pH 6.0.

Lysinyl and amino terminal residues are
reacted with succinic or other carboxylic acid
anhydrides. Derivatization with these agents has the
15   effect or reversing the charge of the lysinyl residues.
Other suitable reagents for derivatizing primary amine
containing residues include imidoesters such as methyl
picolinimidate; pyridoxal phosphate; pyridoxal; chloro-
borohydride; trinitrobenzenesulfonic acid; O-methy-
20   lisourea; 2,4 pentanedione; and transaminase-catalyzed
reaction with glyoxylate.

Arginyl residues are modified by reaction with
one or several conventional reagents, among them phenyl-
glyoxal, 2,3-butanedione, 1,2-cyclohexanedione, and
25   ninhydrin. Derivatization of arginine residues requires
that the reaction be performed in alkaline conditions
because of the high pK_a of the guanidine functional
group. Furthermore, these reagents may react with the

71

groups of lysine as well as the arginine alpha-amino group.

Tyrosyl residues are well-known targets of modification for introduction of spectral labels by reaction with aromatic diazonium compounds or tetranitromethane. Most commonly, N-acetylimidizol and tetranitromethane are used to form O-acetyl tyrosyl species and 3-nitro derivatives, respectively.

Carboxyl side groups (aspartyl or glutamyl) are selectively modified by reaction carbodiimide (R'-N-C-N-R') such as 1-cyclohexyl-3-(2-morpholinyl(4-ethyl) carbodiimide or 1-ethyl-3-(4-azonia-4,4-dimethylpentyl) carbodiimide. Furthermore, aspartyl and glutamyl residue are converted to asparaginyl and glutaminyl residues by reaction with ammonium ions.

Glutaminyl and asparaginyl residues are frequently deamidated to the corresponding glutamyl and aspartyl residues. Alternatively, these residues are deamidated under mildly acidic conditions. Either form of these residues falls within the scope of this invention.

Derivatization with bifunctional agents is useful, for example, for cross-linking the peptide to a water-insoluble support matrix or to other macromolecular carriers. Commonly used cross-linking agents include, for example, 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters

72

such as 3,3'-dithiobis(succinimidylpropionate), and
bifunctional maleimides such as bis-N-maleimido-1,8-
octane. Derivatizing agents such as methyl-3-[p-
azidophenyl) dithiolpropioimidate yield photoactivatable

5   intermediates that are capable of forming crosslinks in
the presence of light. Alternatively, reactive water-
insoluble matrices such as cyanogen bromide-activated
carbohydrates and the reactive substrates described in
U.S. Patent Nos. 3,969,287; 3,691,016; 4,195,128;

10   4,247,642; 4,229,537; and 4,330,440 are employed for
protein immobilization.

Other modifications include hydroxylation of
proline and lysine, phosphorylation of hydroxyl groups
of seryl or threonyl residues, methylation of the alpha-

15   amino groups of lysine, arginine, and histidine side
chains (Creighton, T.E., Proteins: Structure and Mole-
cular Properties, W.H. Freeman & Co., San Francisco, pp.
79-86 (1983)), acetylation of the N-terminal amine, and,
in some instances, amidation of the C-terminal carboxyl

20   groups.

Such derivatized moieties may improve the
stability, solubility, absorption, biological half life,
and the like. The moieties may alternatively eliminate
or attenuate any undesirable side effect of the protein

25   and the like. Moieties capable of mediating such
effects are disclosed, for example, in Remington's
Pharmaceutical Sciences, 18th ed., Mack Publishing Co.,
Easton, PA (1990).

73

The term "fragment" is used to indicate a polypeptide derived from the amino acid sequence of the protein having a length less than the full-length polypeptide from which it has been derived. Such a

5   fragment may, for example, be produced by proteolytic cleavage of the full-length protein. Preferably, the fragment is obtained recombinantly by appropriately modifying the DNA sequence encoding the proteins to delete one or more amino acids at one or more sites of

10 the C-terminus, N-terminus, and/or within the native sequence. It is understood that such fragments, may retain one or more characterizing portions of the native protein. Examples of such retained characteristics include: catalytic activity; substrate specificity;

15 interaction with other molecules in the intact cell; regulatory functions; or binding with an antibody speci-fic for the native protein, or an epitope thereof.

Another functional derivative intended to be within the scope of the present invention is a "variant"

20 polypeptide which either lack one or more amino acids or contain additional or substituted amino acids relative to the native polypeptide. The variant may be derived from a naturally occurring protein by appropriately modifying the protein DNA coding sequence to add,

25 remove, and/or to modify codons for one or more amino acids at one or more sites of the C-terminus, N-terminus, and/or within the native sequence. It is understood that such variants having added, substituted and/or additional amino acids retain one or more

74

characterizing portions of the native protein, as
described above.

A functional derivative of a protein with
deleted, inserted and/or substituted amino acid residues
5   may be prepared using standard techniques well-known to
those of ordinary skill in the art. For example, the
modified proteins may be produced using site-directed
mutagenesis techniques (as exemplified by Adelman
et al., 1983, DNA 2:183) wherein nucleotides in the DNA
10   coding the sequence are modified such that a modified
coding sequence is modified, and thereafter expressing
this recombinant DNA in a prokaryotic or eukaryotic host
cell, using techniques such as those described above.
Alternatively, proteins with amino acid deletions,
15   insertions and/or substitutions may be conveniently
prepared by direct chemical synthesis, using methods
well-known in the art. The functional derivatives
typically exhibit the same qualitative biological
activity as the native proteins.

20   **XVI Administration**

Administration as used herein refers to the
route of introduction of a vector or carrier of DNA into
the body. Administration may include intravenous,
intramuscular, topical, or oral methods of delivery.
25   Administration can be directly to a target tissue or
through systemic delivery.

In particular, the present invention can be
used for treating disease or for administering the

75

formulated DNA expression vectors capable of expressing
any specific nucleic acid sequence. Administration can
also include administering a regulatable vector
discussed above. Such administration of a vector can be
5   used to treat disease. The preferred embodiment is by
direct injection to the target tissue or systemic
administration.

A second critical step is the delivery of the
DNA vector to the nucleus of the target cell where it
10   can express a gene product. In the present invention
this is accomplished by formulation. The formulation
can consist of purified DNA vectors or DNA vectors
associated with other formulation elements such as
lipids, proteins, carbohydrates, synthetic organic or
15   inorganic compounds. Examples of such formulation
elements include, but are not limited to, lipids capable
of forming liposomes, cationic lipids, hydrophilic
polymers, polycations (e.g., protamine, polybrene,
spermidine, polylysine), peptide or synthetic ligands
20   recognizing receptors on the surface of the target
cells, peptide or synthetic ligands capable of inducing
endosomal lysis, peptide or synthetic ligands capable of
targeting materials to the nucleus, gels, slow release
matrices, soluble or insoluble particles, as well as
25   other formulation elements not listed. This includes
formulation elements for enhancing the delivery, uptake,
stability, and/or expression of genetic material into
cells.

76

The delivery and formulation of any selected vector construct will depend on the particular use for the expression vectors. In general, a specific formulation for each vector construct used will focus on
5   vector uptake with regard to the particular targeted tissue, followed by demonstration of efficacy. Uptake studies will include uptake assays to evaluate cellular uptake of the vectors and expression of the tissue specific DNA of choice. Such assays will also determine
10 the localization of the target DNA after uptake, and establishing the requirements for maintenance of steady-state concentrations of expressed protein. Efficacy and cytotoxicity can then be tested. Toxicity will not only include cell viability but also cell function.

15   DNA uptake by cells associated with fluid spaces have the unique ability to take up DNA from the extracellular space after simple injection of purified DNA preparations into the fluid spaces. Expression of DNA by this method can be sustained for several months.

20   Incorporating DNA by formulation into particulate complexes of nanometer size that undergo endocytosis increases the range of cell types that will take up foreign genes from the extracellular space.

            Formulation can also involve DNA transporters
25 which are capable of forming a non-covalent complex with DNA and directing the transport of the DNA through the cell membrane. This may involve the sequence of steps including endocytosis and enhanced endosomal release. It is preferable that the transporter also transport the

77

DNA through the nuclear membrane. See, *e.g.*, the
following applications all of which (including drawings)
are hereby incorporated by reference herein: (1) Woo et
al., U.S. Serial No. 07/855,389, entitled "A DNA
5   Transporter System and Method of Use" filed March 20,
1992; (2) Woo et al., PCT/US93/02725, entitled "A DNA
Transporter System and method of Use", (designating the
U.S. and other countries) filed March 19, 1993; and (3)
continuation-in-part application by Woo et al., entitled
10   "Nucleic Acid Transporter Systems and Methods of Use",
filed December 14, 1993, assigned attorney docket number
205/012 but not yet assigned a U.S. Serial Number.

In addition, delivery can be cell specific or
tissue specific by including cell or tissue specific
15   promoters. Furthermore, mRNA stabilizing sequences (3'
UTR's) can be used to provide stabilized modified
receptor molecules. Such stabilizing sequences increase
the half-life of mRNAs and can be cell or tissue
specific. The above is discussed in more detail in U.S.
20   Patent 5,298,422 (Schwartz et al.) and U.S. Application
Serial Number 08/209,846 (Schwartz et al.), filed March
9, 1994, entitled "Expression Vector Systems and Method
of Use." Both of these, the whole of which, are
incorporated by reference herein, including drawings.
25   Information regarding endothelial specific sequences is
provided in U.S. Patent Application Serial No.
08/146,930, filed Novemeber 1, 1993 and Serial No.
08/147,777, filed Novemeber 1, 1993 both of which are
incorporated herein by reference in their entirety

78

including any drawings. Information regarding muscle
specific sequences is provided in U.S. Patent
Application Serial No. 08/472,809.

In a preferred method of administration
5   involving a DNA transporter system, the DNA transporter
system has a DNA binding complex with a binding molecule
capable of non-covalently binding to DNA which is
covalently linked to a surface ligand. The surface
ligand is capable of binding to a cell surface receptor
10   and stimulating entry into the cell by endocytosis,
pinocytosis, or potocytosis. In addition, a second DNA
binding complex is capable of non-covalently binding to
DNA and is covalently linked to a nuclear ligand. The
nuclear ligand is capable of recognizing and
15   transporting a transporter system through a nuclear
membrane. Additionally, a third DNA binding complex may
be used which is also capable of non-covalently binding
to DNA. The third binding molecule is covalently linked
to an element that induces endosomal lysis or enhanced
20   release of the complex from the endosome after
endocytosis. The binding molecules can be spermine,
spermine derivatives, histones, cationic peptides and/or
polylysine. See also Szoka, C.F., Jr. et al.,
Bioconjug. Chem. 4:85-93 (1993); Szoka, F.C., Jr. et
25   al., P.N.A.S., 90:893-897 (1993).

Transfer of genes directly has been very
effective. Experiments show that administration by
direct injection of DNA into joint tissue results in
expression of the gene in the area of injection.

79

Injection of plasmids containing the mutated receptors
into the spaces of the joints results in expression of
the gene for prolonged periods of time. The injected
DNA appears to persist in an unintegrated

5   extrachromosomal state. This means of transfer is the
preferred embodiment.

The formulation used for delivery may also be
by liposomes or cationic lipids. Liposomes are hollow
spherical vesicles composed of lipids arranged in a

10   similar fashion as those lipids which make up the cell
membrane. They have an internal aqueous space for
entrapping water soluble compounds and range in size
from 0.05 to several microns in diameter. Several
studies have shown that liposomes can deliver nucleic

15   acids to cells and that the nucleic acid remains
biologically active. Cationic lipid formulations such
as formulations incorporating DOTMA has been shown to
deliver DNA expression vectors to cells yielding
production of the corresponding protein. Lipid

20   formulations may be non-toxic and biodegradable in
composition. They display long circulation half-lives
and recognition molecules can be readily attached to
their surface for targeting to tissues. Finally, cost
effective manufacture of liposome-based pharmaceuticals,

25   either in a liquid suspension or lyophilized product,
has demonstrated the viability of this technology as an
acceptable drug delivery system. See Szoka, F.C., Jr.
et al., Pharm. Res., 7:824-834 (1990); Szoka, F.C., Jr.
et al., Pharm. Res., 9:1235-1242 (1992).

80

The chosen method of delivery should result in nuclear or cytoplasmic accumulation and optimal dosing. The dosage will depend upon the disease and the route of administration but should be between 1-1000 µg/kg of

5   body weight. This level is readily determinable by standard methods. It could be more or less depending on the optimal dosing. The duration of treatment will extend through the course of the disease symptoms, possibly continuously. The number of doses will depend

10   upon disease, the formulation and efficacy data from clinical trials.

With respect to vectors, the pharmacological dose of a vector and the level of gene expression in the appropriate cell type includes but is not limited to

15   sufficient protein or RNA to either: (1) increase the level of protein production; (2) decrease or stop the production of a protein; (3) inhibit the action of a protein; (4) inhibit proliferation or accumulation of specific cell types; and (5) induce proliferation or

20   accumulation of specific cell types. As an example, if a protein is being produced which causes the accumulation of inflammatory cells within the joint, the expression of this protein can be inhibited, or the action of this protein can be interfered with, altered,

25   or changed.

EXAMPLES

The examples below are non-limiting and are merely representative of various aspects and features of

81

the present invention. The examples below demonstrate
the isolation, and characterization of SRC-1 and provide
evidence indicating that SRC-1 is a coactivator for
steroid receptors.

5   **EXPERIMENTAL PROCEDURES**

                All reagent for buffers and solutions were
certified A.C.S. grade. When required, molecular grade
reagents were used. All recombinant DNA works were
performed according to standard procedures (Ausubel,

10   F.M., et al., (1992), Short Protocols In Molecular
        Biology. New York: Greene Publishing Associates and
        John Wiley & Sons). Plasmid DNA constructs were
        prepared and purified using the plasmid DNA isolation
        kit (QIAGEN) and confirmed by sequence using deaza-

15   dideoxy NTPs and Sequenase version 2.0 (US.
        biochemical).

Plasmid construction.

                The $GAL4_{DBD}$-PR_{LBD} chimeric protein was
constructed in the pAS1 yeast expression plasmid

20   (Durfee, T. et al., (1993), Genes and Dev. 7, 555-569).
        The coding sequence from amino acid 631-933 of hPR_B in
        the vector $YEphPR_B$ (Vegeto, E. et al., (1992), Cell 69,
        703-713) was PCR amplified using Taq polymerase
        (Promega) with the primers CGCCATGGTCCTTGGAGGT (upper

25   strand) and AAGTCGACACATTCACTTTTTATGAAAGAGAAG (lower
        strand).

82

The primers were designed to contain NcoI and
SalI restriction enzyme sites at 5' and 3' end of the
amplified product. Amplification was carried out
according to manufacturer conditions for 40 cycles (2
5   min 94°C, 2 min 45°C, 3 min 72°C) with a final extension
of 10 min at 72°C. The amplified product was then
double-digested with NcoI and SalI, gel purified and
inserted into the NcoI-SalI site of the vector pAS1
(Durfee, T. et al., (1993), Genes and Dev. 7, 555-569).
10   For the two-hybrid screening, the NcoI-SalI insert was
transferred into the pAS1-cyh vector which contains the
cycloheximide marker for curing and selection in the
Y190 yeast strain. The correct expression of the
GAL4$_{DBD}$-PRL$_{BD}$ fusion protein in yeast was assessed by
15   Dot and Western immunoblotting using the C262 monoclonal
antibody and by hormone binding assays as previously
described (Weigel, N.L. et al., (1992), Mol. Endocrinol.
6, 1585-1597).

For in vitro transcription and translation,
20   SRC-1(.8) cDNA XhoI fragment was cloned into the SalI
site of the pT7BSalI vector (Baniahmad, A. et al.,
(1993), Proc. Natl. Acad. Sce. USA 90, 8832-8836). The
SRC-1mut was constructed by cloning the SRC-1(.8) BamHI-
BglII fragment into the BglII-BamHI sites of the pABWgal
25   mammalian expression vector (Baniahmad, A. et al.,
(1993), Proc. Natl. Acad. Sci. USA 90, 8832-8836).

The mammalian expression vectors and their
reporter plasmids for hPR$_{B}$ (Vegeto, E. et al., (1992),
Cell 69, 703-713) and PRE2-TATA-CAT (Beekman, J.M. et

83

al., (1993), <u>Mol. Endocrinol.</u> 7, 1266-1274), hER (Smith,
C.L. et al., (1993), <u>Proc. Natl. Acad. Sci. USA</u> *90*,
6120-6124) and ERE2-TATA-CAT (Beekman, J.M. et al.,
(1993), <u>Mol. Endocrinol.</u> 7, 1266-1274), hGR (Giguere, V.

5   et al., (1986), <u>Cell</u> *46*, 645-652), hTR (Umesono, K. et
al., (1991), <u>Cell</u> *65*, 1255-12661), mouse RXR (Leng et
al., 1994), DR4-tk-CAT and DR1-tk-CAT (Cooney, A.J. et
al., (1993), <u>J. Biol. Chem.</u> *268*, 4152-4160), E2F and
E2F-tk-CAT (Helin, K. et al., (1993), <u>Mol. Cell. Biol.</u>

10   *13*, 6501-6508), Sp1 and Sp1-tk-CAT (Courey, A.J. and
Tjian, R. (1988), <u>Cell</u> *55*, 887-898), CREB and CRE-tk-CAT
(Chrivia, J.C. et al., (1993), <u>Nature</u> *365*, 855) and
GAL4VP16 and 17mer-tk-CAT (Baniahmad, A. et al., (1992),
<u>EMBO J.</u> *11*, 1015-1023) have been previously described

15   and the above references are incorporated herein by
reference in their entirety, including any drawings.

Two-hybrid screening.

The yeast strain Y190 containing the GAL4$_{DBD}$-
PR$_{LBD}$ expression plasmid was transformed with a human B-

20   lymphocyte cDNA expression library constructed in the
yeast expression vector pACT and the transformant
screened for interacting proteins in the presence of 10^{-6}
M progesterone as described by Durfee, T. et al.,
(1993), <u>Genes and Dev.</u> 7, 555-569. Library cDNA

25   plasmids from positives clones were recovered and used
to retransform Y190 cells containing the GAL4$_{DBD}$-PR$_{LBD}$
or empty expression vectors as indicated herein.

84

The --galactosidase activities of the
transformants in liquid culture were determined as
described (Ausubel, F.M., et al., (1992), Short
Protocols In Molecular Biology. New York: Greene
5   Publishing Associates and John Wiley & Sons) using O-
Nitrophenyl --D-galactopyranoside (ONPG) as a substrate.
The specificity of the interacting proteins was assessed
by mating Y190 cells containing the SRC-1(.8) cDNA with
the Y187 strain containing pAS1-SNF, pAS1-p53, pAS1-CDK
10   or pAS1-lamin and the --galactosidase activity of the
diploids were determined by filter lift or in liquid
culture assays (Ausubel, F.M., et al., (1992), Short
Protocols In Molecular Biology, New York: Greene
Publishing Associates and John Wiley & Sons; Durfee, T.
15   et al., (1993), Genes and Dev. 7, 555-569).

In vitro protein-protein interactions.

Receptor-specific affinity resins were
constructed by linking recombinant baculovirus
glutathione-S-transferase hPR$_A$ (GST-hPR$_A$) to GST-
20   Sepharose beads as previously described (Baniahmad, A.
et al., (1995), Mol. Cell Biol. 15, 76-86). The GST-PR$_A$
fusion protein was constructed by inserting the PR$_A$ cDNA
into baculovirus expression vector according to standard
procedures (Beekman, J.M. et al., (1994), Gene 146, 285-
25   289). Receptors were activated in vivo by the addition
of 10^{-8} M hormones (progesterone or RU486) to intact
cells for 24 h before harvesting. Receptors were then
prepared as whole-cell extracts and treated for an

85

additional 15 min at 30°C with 10^{-6} M hormone before
purification. Approximately 400 μg of total proteins
were incubated with 20 μl of GST-Sepharose beads in
suspension (Pharmacia) for 2 h at 4°C. Resins were

5   then washed twice with NENT buffer (20 mM Tris-OH pH 8.0
containing 100mM NaCl, 1mM EDTA, 0.5% NP40 and 0.5% milk
powder) and twice more with transcription buffer (20 mM
HEPES pH 7.9 containing 60 mM NaCl, 1mM dithiothreitol,
6mM MgCl2, 0.1 mM EDTA and 10% Glycerol).

10   Subsequently, beads containing purified
receptors were mixed with 25 μl of crude lysate of in
vitro transcribed and translated [^{35}S]met-SRC-1(.8)
(Promega) and interactions were allowed to occur at 4°C
for 1 h in an end-over-end rotator in 200 μl of

15   transcription buffer. After interactions, beads were
washed once with NENT buffer, three times with NENT
buffer without milk powder and twice with transcription
buffer. Bound proteins were then eluted with 0.2%
sodium dodecyl sulfate (SDS) in 10 mM Tris-OH buffer

20   (pH7.6), fractionated on SDS-PAGE and subjected to
fluorography for ^{35}S as described (Baniahmad, C. et al.,
(1995), Mol. Endocrinol. 9, 34-43). The input lane
represents 10% of the total volume of the crude lysate
used in each reaction. The amount of receptor in the

25   beads was estimated by commassie blue-staining of the
GST-hPR$_A$ eluted from the affinity resins and
fractionated on SDS-PAGE.

86

RNA analysis.

Northern blot membranes containing poly(A)⁺
RNAs from different human organs (Clontech) or isolated
from different cell lines (Invitrogen) were hybridized
5   at 42°C in 50% formamide with 5x10⁵ cpm/ml of the [³²P]
random primed labeled SRC-1(.8) XhoI insert for 36 h as
described (Ausubel, F.M., et al., (1992), Short
Protocols In Molecular Biology, New York: Greene
Publishing Associates and John Wiley and Sons). After
10   hybridization, a final wash in 0.25 X SSC, 0.1% SDS at
65°C was performed and then membranes submitted to
autoradiography using Kodak X-OMAT Imaging films.

Cloning of full-length cDNA encoding SRC-1.

Following the identification of SRC-1(.8), the
15   XhoI insert cDNA was used to screen a fibroblast library
constructed in 6ZAP (Pereira, F. et al., (1992),
Biochem. Biophys. Res. Commun. 175, 831-838) according
to standard procedures (Ausubel, F.M., et al., (1992),
Short Protocols In Molecular Biology, New York: Greene
20   Publishing Associates and John Wiley & Sons). Several
clones were identified and isolated. Sequence
comparison revealed that two of them, 154-22a of 1.4kb
and 154-25 of 2.3kb, encompass ~3.6 kb of the most 3'
end of SRC-1. The clone 154-22a also contained
25   additional 0.8 kb of 5' sequence not related to SRC-1,
which were excised by subcloning the XhoI (filled)
insert into the PvuII site of pABWgal vector. The
remaining 5' end sequence of SRC-1 was cloned by PCR

87

amplification from a 6gt11 HeLa library (Clontech) using
10[9] phages with the 6gt11 forward primer and the nested
primer GGAATTCCCGACGTTGTGCCAACA.

5   Amplification was carried out using Taq
polymerase under manufacturer conditions in two steps.
First, amplification was performed for 1 min 94°C, 1 min
72°C, 2 min 72°C followed for another five cycles with a
progressive decrease in the annealing temperature of one
degree Celsius per cycle, from 71°C to 67°C. Then,
10   amplification continued for 29 cycles (1 min 94°C, 1 min
64°C, 2 min 72°C) with a final extension of 5 min 72°C.
The PCR products were cloned into the pCRII TA-cloning
vector (Invitrogen) using either the direct PCR product
or the >1.6 kb sized product purified on a 1% agarose
15   gel. Clones ranging from 0.6 to 2.2 kb were subjected
to sequence using the nested oligo as primer. Amplified
cDNAs containing identical sequence to the 5' end of
154-22a were assigned as positives. The EcoRI (partial)
and BsmI insert from the longest cDNA amplified (2.2 kb)
20   was ligated to the BsmI-SalI insert from 154-22a and
then religated into the EcoRI-XhoI sites of the
mammalian expression vector pBK-CMV (Strategene) and
renamed SRC-1. The sequence of the PCR amplified open
reading frame of SRC-1 represents the data originated
25   from two independent PCR-amplified products. The 3' end
untranslated region of SRC-1 from clone 154-25 (2.3 kb)
was not included.

88

Transient transfection and CAT assays.

        Cells were maintained in Dulbecco's modified
Eagle's medium supplemented with 10% (vol/vol) fetal
calf serum (Smith, C.L. et al., (1993), <u>Proc. Natl.</u>
5   <u>Acad. Sci. USA</u> *90*, 6120-6124). The day before
transfection, 10^6 cells were seeded in 100 mm dishes and
4-6 h later switched serum-free medium supplemented with
Nutridoma-SR (Boehringer Mannhein). When ER was used
for transfections, phenol-red-free medium was used
10 (Smith, C.L. et al., (1993), <u>Proc. Natl. Acad. Sci. USA</u>
90, 6120-6124). HeLa cells were transfected with
mammalian expression vectors encoding steroid receptors
or other transcription factors along with 5 µg of CAT
reporter. Lipofectin (GibcoBRL) was selected for
15 transfection of Hela cells. CV1 and Lmtk- cells were
transfected using polybrene (Sigma) as previously
described (Denner, L.A. et al., (1990), <u>Science</u> *250*,
1740-1743).

        After transfection, excess DNA was removed and
20 cells were treated for 42 h with media containing
hormones as indicated in figure legends. Cells were
then harvested and proteins prepared as whole-cell
extracts by freeze-thaw lysis. The reporter CAT
activity in the extract was determined using 100 uCi of
25 [^{14}C]-chloramphenicol and 4mM acetyl coenzyme A as
substrate and standardized by protein content. Activity
of the extract was calculated by determining the
percentage of conversion of [^{14}C]-chloramphenicol to the
mono and diacetylated forms (Ausubel, F.M., et al.,

89

(1992), <u>Short Protocols In Molecular Biology.</u> New York:

Greene Publishing Associates and John Wiley & Sons).

Example 1:

Isolation and characterization of PR interacting

5   **proteins.**

Human PR contains a ligand-inducible

transactivation function in the C-terminus of its

ligand-binding domain (Meyer, M.E. et al., (1990), <u>EMBO</u>

<u>J.</u> 9, 3923-3932; Gronemeyer, H. (1991), <u>Ann. Rev. Genet.</u>

10   25, 89-123). To isolate cDNAs encoding proteins that

specifically interact with the hormone binding domain of

PR, we used the yeast two-hybrid system previously

described by Durfee, T. et al., (1993), <u>Genes and Dev.</u>

7, 555-569, incorporated herein by reference in its

15   entirety, including any drawings. First we constructed

a chimeric protein between the yeast transcription

factor GAL4$_{DBD}$ (amino acid 1-147) and the region of hPR

that encompasses the hinge and the LBD (PR$_{LBD}$, amino

acid 362-933) in the vector pAS1. Yeast Y190 cells

20   expressing GAL4$_{DBD}$-PR$_{LBD}$ fusion protein were then

transformed with a B lymphocyte cDNA library fused to

the yeast GAL4 activation domain (GAL4$_{AD}$).

The yeast strain Y190 provides a dual reporter

system to screen for cDNAs encoding proteins that

25   interact with PR$_{LBD}$: *HIS3* for histidine prototrophy and

LacZ for --galactosidase activity, both of which are

chromosomally integrated and their expression regulated

by a GAL4 promoter. Interaction between GAL4$_{DBD}$-PR$_{LBD}$

and the cDNA encoded protein fused to the GAL4$_{AD}$ (amino
acid 768 to 880) will result in the activation of the
chromosomal integrated HIS3 and LacZ gene transcription
under the control of the galactose-inducible UAS$_G$

5   promoter. Cells grown in selective media for histidine
prototrophy and for their blue phenotype in the presence
of X-gal and 10^{-6} M progesterone. From approximately
600,000 yeast colonies that were screened in the
presence of 10^{-6} M progesterone, seven positive clones

10   were obtained. Their interaction with the PR$_{LBD}$ was
specific since they failed to interact with other
unrelated proteins such as p53, lamin, CDK and SNF1,
when fused to the GAL4$_{DBD}$.

        From those seven isolated cDNAs, the one [SRC-

15   1(.8)] that exhibited the strongest interaction with the
PR$_{LBD}$ in the presence of 10^{-6} M progesterone, was
selected for further studies. Galactosidase activity
was only observed when the PR$_{LBD}$ was coexpressed with
SRC-1(.8) cDNA fused to the GAL4 activation domain.

20   Neither SRC-1(.8) nor PR$_{LBD}$ fusion proteins were active
when expressed alone. SNF1 and SNF4, two proteins
previously described to interact in the two-hybrid
system were used as a positive control (Durfee et al.,
(1993), Genes and Dev. 7:555-569). In order to

25   eliminate the possibility that failure of the GAL4$_{DBD}$-
PR$_{LBD}$ construct to activate the LacZ reporter is due to
lack of expression, the GAL4$_{DBD}$-PR$_{LBD}$ fusion protein
level was determined. Dot- and Western immunoblotting
which were carried out with the C262 anti-hPR monoclonal

91

antibody that recognizes the last 14 amino acid in the
C-terminus of hPR (Weigel, N.L. et al., (1992), Mol.
Endocrinol. 6, 1585-1597) indicates that the GAL4$_{DBD}$-
PR$_{LBD}$ fusion protein was expressed correctly in yeast
5   cells. In addition, ligand-binding assays revealed that
the GAL4$_{DBD}$-PR$_{LBD}$ fusion protein was expressed in yeast
at 0.93 ± 0.12 (Xn=7 ± SE) pmol/mg of protein.

Example 2:

SRC-1 Interacts with receptor in a ligand-dependent
10   **manner.**

        We examined whether the interaction of SRC-
1(.8) with PR was hormone dependent. First, we carried
out interactions in the yeast two-hybrid system.   Yeast
cells expressing GAL4$_{DBD}$ - PR$_{LBD}$ and SRC-1(.8) were
15   grown in selective media containing 10^{-6} M progesterone,
no hormone or 10^{-6} M RU486 and the --galactosidase
activity determined as in Example 1. The interaction
was significantly reduced when the receptor was either
free of ligand or bound to RU486. Both progesterone and
20   RU486 had no effect on the activity of PR$_{LBD}$ or SRC-
1(.8) alone; no major effect of the ligands on the
SNF1/SNF4 positive control was observed.   To
substantiate further that the ligand-dependent
interaction observed in intact cells is due to a direct
25   interaction with the receptor protein, we also performed
an in vitro binding assay using PR$_A$. The hPR$_A$ cDNA was
fused to glutathione-S-transferase (GST) sequence in a
baculovirus expression vector to generate GST-PR$_A$ fusion

92

protein suitable for expression in Sf9 insect cells as
described herein, i.e. radiolabeled [^{35}S] met SRC-1(.8)
was incubated in batch with purified baculovirus
expressed GST-PR$_A$ fusion protein bound to glutathione-

5   Sepharose beads either in the absence or in the presence
of 10^{-6} M progesterone or RU486. GST-PR$_A$ was linked to
glutathione sepharose beads and used as an affinity
matrix for binding studies. Bound radiolabeled [35S]
met SRC-1(.8) was then eluted and analyzed in a 15% SDS-

10   PAGE and subjected to fluorography as described herein.
SDS-PAGE analysis of the GST-PR$_A$ eluted from the
glutathione affinity column indicated that it is indeed
expressed in a full-length form.

In vitro transcribed and translated [^{35}S]SRC-

15   1(.8) was retained by the GST-PR$_A$ affinity columns when
the receptor was bound to progesterone. Significantly
lesser interactions were observed with ligand-free or
RU486 bound receptor. Little or no binding was observed
on the GST protein column without PR$_A$.

20   Thus, interaction of SRC-1(.8) with PR
occurred in vivo and in vitro in a ligand-dependent
manner and this interaction involved the ligand binding
domain. Our conclusion was substantiated further by an
observation that the N-terminal region of PR$_A$ (amino

25   acids 165-565) fused to the GAL4$_{DBD}$ failed to interact
with SRC-1(.8) in the two-hybrid system.

93

Example 3:

Isolation of full-length SRC-1.

         To obtain a full-length cDNA, conventional
screening of a fibroblast library with SRC-1(.8 kb) cDNA
5   as probe was carried out. We isolated two clones of 1.4
kb and 2.3 kb that encompass most of the 3' end of SRC-
1. The 5' end of SRC-1 (2.2 kb), was cloned by PCR from
a HeLa cell library using nested primers as described
herein. Sequence alignment of these three isolated
10   clones revealed a 5.6 kb cDNA containing an open reading
frame of 1061 amino acids with a predicted mass of 114.1
kDa and apparent molecular size of ~125 kDa on SDS-PAGE
after in vitro transcription and translation.

         Current sequence comparisons using the BLAST
15   algorithm show that amino acids 605 to 1005 are
identical to the hin-2 gene. This gene was identified
by analysis of a human immunodeficiency virus type 1
promoter insertion in vivo (access number U19179).
Sequence comparison reveals that the hin-2 gene contains
20   a 54 bp insertion. This DNA insertion introduces a stop
codon after 14 amino acids downstream, resulting in a
premature termination of SRC-1. In addition, partial
DNA sequences for SRC-1 have been isolated randomly
(access numbers T56159 and U19179). No function for
25   these partial cDNAs has been described. SRC-1 sequence
comparisons at the amino acid level show no significant
homology to any known protein. A notable glutamine rich
region (31.4% Q) is observed between residues 673 and
758. The N-terminus of SRC-1 (residues 258-350) is

94

serine and threonine rich (22.6% S and 11.8% T). In
addition, there are regions noted for their leucine and
proline content (9.2% L and 9.6% P).

Example 4:

5   **Expression of SRC-1 mRNAs.**

Northern blot analysis using poly A⁺ RNA from
human tissue and cell lines was carried out to determine
the size and distribution of SRC-1 mRNA. Hybridization
was performed with the [^{32}P] radiolabeled SRC-1(.8) cDNA.

10   Analysis of actin mRNA was performed to control for
loading and integrity of the RNAs. SRC-1 is expressed
as two mRNAs of approximately 5.5 and 7.5 kb in a
variety of tissues, including heart, placenta, lung,
liver, smooth muscle, kidney and pancreas. In brain,

15   SRC-1 was predominantly expressed as a 7.5 kb mRNA. All
other human tissues analyzed (spleen, thymus, prostate,
testis, ovary, small intestine, colon and leukocytes),
as well as cell lines (HeLa, CV-1 and Jurkat) exhibited
the same two SRC-1 mRNAs, but with some variation in

20   expression level.

Example 5:

**Transient expression of SRC-1 stimulates ligand-bound
hPR transactivation.**

To further investigate the role of SRC-1 in

25   receptor transactivation of target gene expression,
transient transfection assays were carried out in
mammalian cell lines. HeLa cells were cotransfected

95

with mammalian expression vectors for full-length SRC-1 and hPR$_B$ together with a plasmid containing two copies of the progesterone-responsive DNA element (PRE) inserted in front of the TATA box of the E1b adenovirus,

5   linked to the chloramphenicol acetyl transferase (CAT) reporter gene, PRE2-TATA-CAT.

In particular, HeLa cells were transiently transfected with 5 μg of PRE2-TATA-CAT or ERE2-TATA-CAT reporter plasmids along with 0.5 μg of hPR$_B$ mammalian

10 expression plasmid and 3 μg of SRC-1 or empty expression vector, either in the absence or in the presence of 2×10^{-8} M progesterone agonist R5020. Cells were harvested 42 h after transfection and the chloramphenicol acetyl transferase (CAT) activity determined using 8 μg of

15 protein extract.

Ligand-free hPR has minimal activity on the reporter either in the absence or in the presence of SRC-1. Addition of the progesterone agonist R5020 resulted in a ~5 fold induction of hPR activity.

20 However, when SRC-1 full-length cDNA was coexpressed with hPR in the presence of hormone, a 14.0 ± 3.3 (Xn =7± SEM) fold induction over the hormone-induced receptor activity was observed (an average of seven experiments).

25 To further confirm that the stimulation observed by coexpression of SRC-1 with hPR was solely dependent on hPR, and not to an indirect effect on the basal activity of the reporter, we replaced the PRE reporter construct with an ERE reporter, ERE2-TATA-CAT.

96

In particular, HeLa cells were transfected with PRE2-TATA-CAT and hPR$_B$ along with SRC-1 or empty expression vector in the absence or in the presence of $2\times10^{-8}M$ R5020 alone or in combination with $4\times10^{-8}M$ RU486

5   and then assayed for CAT activity using 14 μg of proteins extract.

In the absence of PR binding site there is no effect of SRC-1 on the reporter construct activity, either in the absence or in the presence of the ligand.

10   These results indicate that SRC-1 enhancement of hPR transactivation occurs via ligand-bound receptor; it does not independently affect the basal transcription machinery. We investigated the effect of SRC-1 on transactivation of hPR bound to the antagonist RU486.

15   Again, SRC-1 enhanced hPR activity in the presence of the agonist R5020. Addition of RU486 to intact cells prevented the hormone-induced transactivation of hPR. Coexpression of SRC-1 with the antagonist-bound receptor was unable to augment and/or recover to any extent the

20   RU486-antagonized receptor transcriptional activity of the reporter gene.

Taken together, these findings indicate that SRC-1 acts directly on the agonist-bound receptor protein to modulate its transcriptional activity without

25   a major effect on the basal promoter and fulfills the general definition of a coactivator. In addition, evidence is provided that a steroid receptor coactivator can discriminate between an agonist and an antagonist receptor complex bound to DNA in vivo.

97

Example 6:

SRC-1 acts as a coactivator for hPR by reversing receptor squelching.

To substantiate the coactivator function of
5   SRC-1 we utilized a more stringent assay for
coactivators that has been proposed to exploit the
property of excess activator to sequester a limited pool
of nuclear coactivator (Flanagan, P.M. et al., (1991),
Nature 350, 436-438). It has been reported earlier that
10 overexpression of ER can squelch PR-dependent
transcriptional activity in cells (Meyer, M.E. et al.,
(1989), Cell 57, 433-442; Conneely, O.M. et al., (1989),
"Promoter specific activating domains of the chicken
progesterone receptor." In Gene Regulation by Steroid
15 Hormones IV. A.K. Roy and J. Clark, eds. (New York,
Berlin, Heidelberg, London, Paris, Tokyo: Springer-
Verlag), pp. 220-223). This finding suggests the
existence of a common limiting factor(s) necessary for
these two receptors to transactivate target genes. We
20 asked whether the coexpression of SRC-1 was able to
reverse this squelching.

HeLa cells were transfected with 1 μg of both
hPR and hER expression vectors, together with 5 μg of
PRE2-TATA-CAT and increasing amounts of SRC-1. Cells
25 were then exposed to ligands R5020 and/or E2 (2 x 10^{-8} M)
and CAT activities determined 42 h later using 40 μg of
proteins extract. The percentage of chloramphenicol
conversion to the mono and diacetylated form was

98

determined and represents the average of three
independent experiments (Xn=3 ± SEM).

The hormone-induced transcriptional activity
mediated by hPR is inhibited ~19 fold upon coexpression

5   of ligand-bound hER. Addition of SRC-1 reversed this
squelching by ~16 fold in a dose-dependent manner.
Could this increase be explained by the enhancement of
residual PR that remains unsquelched during the assay?
We think not since a ~5 fold stimulation of PR

10   transactivation is observed at the highest concentration
of SRC-1 (2 μg), as compared to a 16 fold recovery in
the presence of estrogen. Such quantitative differences
in activation argue strongly for reversal of squelching.
We conclude that SRC-1 is a limiting factor necessary

15   for efficient PR and ER transactivation.

Example 7:
SRC-1 is a coactivator for multiple steroid receptor
superfamily members.

The ability of SRC-1 to reverse squelching

20   between two different members of the steroid receptor
superfamily suggested that SRC-1 might be a general
coactivator for members of the steroid receptor
superfamily. In fact, the N-terminally truncated form
of SRC-1, SRC-1(.8) also was able to interact with

25   baculovirus expressed GST-ER and GST-TR in vitro. Thus,
we investigated the effect of SRC-1 expression on the
transcriptional activity of other intracellular
receptors.

99

HeLa cells were cotransfected with 0.5 μg of
cDNAs encoding various steroid receptors indicated and 5
μg of their cognates HREs containing reporters,
including PR/PRE2-TATA-CAT, ER/ERE2-TATA-CAT, GR/PRE2-
5   TATA-CAT, TR-/DR4-tk-CAT and RXR/DR1-tk-CAT, along with
3 μg of SRC-1 or empty expression vector. Cells were
then treated with their corresponding ligands (PR:R5020,
ER:estradiol, GR:dexamethasone, TR:Triac, RXR:9-cis
retinoic acid) at a concentration of 2×10^{-8}M. The CAT
10   activity of the reporter was determined using 14 μg of
protein extract.

SRC-1 can enhance PR, as well as ER, GR, TR
and RXR transcriptional activity through their cognate
DNA response-elements. Therefore, SRC-1 appears to be a
15   general coactivator for all members of the nuclear
receptors superfamily tested to date.

HeLa cells were transfected as described above
with various activators and their cognates DNA-elements
containing reporter, including the chimera GAL4-
20   VP16/UAS$_G$-tk-CAT, Sp1/Sp1-tk-CAT, E2F/E2F-tk-CAT, and
CREB/CRE-tk-CAT in the absence and in the presence of 1
μM forskolin (Fsk) along with SRC-1 or empty expression
vector. The CAT activity of the reporter was
determined. Each assay is representative of at least
25   two independent experiments.

To determine the spectrum of action of SRC-1
we examined its effect on selected transactivators. We
observed that SRC-1 enhanced the transcriptional
activity of GAL4-VP16 chimera protein through the UAS$_G$

100

element. Albeit to a lesser extent, SRC-1 can also
enhance Sp1 transcriptional activity. In contrast, SRC-
1 did not alter the transcriptional activity of other
nuclear factors such as E2F and E47. We tested whether
5   SRC-1 could influence the transcriptional activity of an
unrelated inducible transcriptional factor such as CREB
(Chrivia, J.C. et al., (1993), Nature 365, 855).

         CREB activity can be induced by the addition
of forskolin to intact cells. Coexpression of SRC-1 did
10   not affect either basal nor forskolin-stimulated
transcriptional activities of CREB. Therefore, while
SRC-1 enhances the activity of members of the
steroid/thyroid superfamily of receptors, it is not a
general coactivator for all classes of transactivators.

15   **Example 8:**
**The truncated C-terminal region of SRC-1 acts as a
dominant negative regulator of steroid receptor
function.**

         To further elucidate the mechanism of SRC-1
20   action, the effect of the N-terminal deletion mutant of
SRC-1, SRC-1(.8), on receptor transactivation was
examined. The amino acid sequence from 865-1061, which
contains the region that binds to the receptor, was
cloned in frame to a mammalian expression vector that
25   provided an exogenous AUG codon.

         Lmtk⁻ cells were transfected with 5 μg of
steroid receptors cDNAs expression vectors and their
reporters, PR/PRE2-tk-CAT and TR$\S$/DR4-tk-CAT, along with

101

10 μg of SRC-1(.8) or empty expression vector and then treated with hormones (2x10⁻⁸ M). The reporter CAT activity was determined using 60 μg of protein extract.

5   The hormone induced transcriptional activity of hPR can be efficiently reduced by the coexpression of SRC-1(.8) in Lmtk- cells. SRC-1(.8) also inhibits the ligand-induced transcriptional activity of TR. No major effect on the hormone-free receptor was observed. SRC-1(.8) also interfered with hPR transcriptional activity 10   in HeLa and CV1 cells to a similar extent. The capability of the truncated SRC-1(.8) to act as dominant-negative repressor on steroid induced receptor transcriptional activity further suggests that SRC-1 is a genuine coactivator for steroid receptor target gene 15   expression. Also these results provide preliminary insight into the functional regions of SRC-1 with respect to transcriptional activation.

          Although certain embodiments and examples have been used to describe the present invention, it will be 20   apparent to those skilled in the art that changes to the embodiments and examples shown may be made without departing from the scope or spirit of the invention.

          Those references not previously incorporated herein by reference, including both patent and non- 25   patent references, are expressly incorporated herein by reference for all purposes.

          Other embodiments are within the following claims.

102

What is claimed is:

CLAIMS

1.   An isolated, enriched or purified nucleic acid encoding a SRC-1 polypeptide.

5   2.   The nucleic acid of claim 1 wherein said nucleic acid is human nucleic acid.

3.   The nucleic acid of claim 1 wherein said nucleic acid encodes at least 25 contiguous amino acids of the amino acid sequence shown in Fig. 1.

10   4.   The nucleic acid of claim 1 wherein said nucleic acid encodes at least 50 contiguous amino acids of the amino acid sequence shown in Fig. 1.

5.   The nucleic acid of claim 1 wherein said nucleic acid encodes at least 100 contiguous amino acids 15 of the amino acid sequence shown in Fig. 1.

6.   The nucleic acid of claim 1 wherein said nucleic acid encodes at least 425 contiguous amino acids of the amino acid sequence shown in Fig. 1.

7.   A nucleic acid probe for the detection of 20 a SRC-1 polypeptide in a sample.

103

8.   The probe of claim 7 wherein said SRC-1
polypeptide comprises at least 25 contiguous amino acids
of the amino acid sequence shown in Fig. 1.

9.   The probe of claim 7 wherein said SRC-1
5 polypeptide comprises at least 50 contiguous amino acids
of the amino acid sequence shown in Fig. 1.

10.   The probe of claim 7 wherein said SRC-1
polypeptide comprises at least 100 contiguous amino
acids of the amino acid sequence shown in Fig. 1.

10   11.   The probe of claim 7 wherein said SRC-1
polypeptide comprises at least 425 contiguous amino
acids of the amino acid sequence shown in Fig. 1.

12.   A composition of matter comprising
recombinant nucleic acid encoding a SRC-1 polypeptide
15 and a vector or a promoter effective to initiate
transcription in a host cell.

13.   The composition of claim 12 wherein said
recombinant nucleic acid encodes at least 25 contiguous
amino acids of the amino acid sequence set forth in Fig.
20 1.

14.   The composition of claim 12 wherein said
recombinant nucleic acid encodes at least 50 contiguous
amino acids of the amino acid sequence set forth in Fig.
1.

104

15.   The composition of claim 12 wherein said
recombinant nucleic acid encodes at least 100 contiguous
amino acids of the amino acid sequence set forth in Fig.
1.

5   16.   The composition of claim 12 wherein said
recombinant nucleic acid encodes at least 425 contiguous
amino acids of the amino acid sequence set forth in Fig.
1.

           17.   A non-human transgenic mammal comprising
10   a transgene encoding a SRC-1 polypeptide.

           18.   The transgenic mammal of claim 17 wherein
said trans gene encodes at least 25 contiguous amino
acids of the amino acid sequence set forth in Fig. 1.

           19.   The transgenic mammal of claim 17 wherein
15   said trans gene encodes at least 50 contiguous amino
acids of the amino acid sequence set forth in Fig. 1.

           20.   The transgenic mammal of claim 17 wherein
said trans gene encodes at least 100 contiguous amino
acids of the amino acid sequence set forth in Fig. 1.

20   21.   The transgenic mammal of claim 17 wherein
said trans gene encodes at least 425 contiguous amino
acids of the amino acid sequence set forth in Fig. 1.

105

22.   A method for increasing the transcription
of a target gene comprising the step of providing
nucleic acid encoding a SRC-1 polypeptide to a cell
containing said target gene.

5   23.   The method of claim 22 further comprising
the step of providing a molecular switch for regulating
expression of a nucleic acid cassette in gene therapy to
said cell containing said target gene,
               wherein said molecular switch comprises a
10   modified steroid receptor, said receptor including a
natural steroid receptor DNA binding domain linked to a
modified ligand binding domain.

24.   The method of claim 23, wherein said
switch is tissue specific.

15   25.   The method of claim 23 further comprising
the step of attaching the molecular switch to a nucleic
acid cassette to form a nucleic acid cassette/molecular
switch complex for use in the gene therapy.

26.   The method of claim 25 further comprising
20   the step of administering a pharmacological dose of the
nucleic acid cassette/molecular switch complex to an
animal or human to be treated.

27.   The method of claim 26, wherein said
molecular switch is turned on or off by dosing the

106

animal or human with a pharmacological dose of a ligand
which binds to the modified ligand binding site.

28.   The method of claim 27, wherein the
nucleic acid sequence is transcribed to produce a
5   protein after the animal or human is given a
pharmacological dose of the ligand.

29.   The method of claim 28, wherein the
molecular switch and the nucleic acid cassette are on
separate plasmids and are co-injected into a target
10   cell.

30.   The method of claim 29 wherein the
molecular switch and the nucleic acid encoding the SRC-1
polypeptide are on separate plasmids and are co-injected
into a target cell.

15   28.   A composition of matter comprising a
molecular switch linked to a nucleic acid cassette,
        wherein said cassette/molecular switch complex
is positionally and sequentially oriented in a vector
such that the nucleic acid in the cassette can be
20   transcribed and when necessary translated in a target
cell; and
        wherein said molecular switch regulates a
constitutively active promoter in a plasmid with a
coding region for a SRC-1 polypeptide.

107

29.   A method for decreasing the transcription
of a target gene comprising the step of providing
nucleic acid encoding a dominant-negative inhibitor of a
SRC-1 polypeptide in a cell containing said target gene.

5   30.   A molecular switch for regulating
expression of a nucleic acid cassette in gene therapy,
comprising:

a modified SRC-1 polypeptide, said polypeptide
including a natural SRC-1 DNA binding domain linked to a
10 modified binding domain.

31.   A method for regulating expression of a
nucleic acid cassette in gene therapy comprising the
step of attaching the molecular switch of claim 30 to a
nucleic acid cassette to form a nucleic acid/molecular
15 switch complex for use in gene therapy and administering
a pharmacological dose of the nucleic acid
cassette/molecular switch complex to an animal or human
to be treated.

32.   A composition of matter comprising a
20 molecular switch linked to a nucleic acid cassette,
wherein said complex is positionally oriented in a
vector such that the nucleic acid in the cassette can be
transcribed and when necessary translated in a target
cell.

108

33.   A method of treating a SRC-1 related
disease or condition comprising the steps of inserting
an expression vector containing a SRC-1 coding sequence
into cells, growing the cells in vitro, and infusing the
5   cells into a patient in need of such treatment.

1/3

```
-56                    ATATCATCGACAGGGAGCACAGTGGGCTTTCTCCTCAAGATGACACTAATTCTGGA  -1

     1 ATG TCA ATT CCC CGA GTA AAT CCC TCG GTC AAT CCT AGT ATC TCT CCA GCT CAT GGT GTG  60
   (1) M   S   I   P   R   V   N   P   S   V   N   P   S   I   S   P   A   H   G   V   (20)

    61 GCT CGT TCA TCC ACA TTG CCA CCA TCC AAC AGC AAC ATG GTA TCC ACC AGA ATA AAC CGC  120
  (21) A   R   S   S   T   L   P   P   S   N   S   N   M   V   S   T   R   I   N   R   (40)

   121 CAG CAG AGC TCA GAC CTT CAT AGC AGC AGT CAT AGT AAT TCT AGC AAC AGC CAA GGA AGT  180
  (41) Q   Q   S   S   D   L   H   S   S   S   H   S   N   S   S   N   S   Q   G   S   (60)

   181 TTC GGA TGC TCA CCC GGA AGT CAG ATT GTA GCC AAT GTT GCC TTA AAC AAA GGA CAG GCC  240
  (61) F   G   C   S   P   G   S   Q   I   V   A   N   V   A   L   N   K   G   Q   A   (80)

   241 AGT TCA CAG AGC AGT AAA CCC TCT TTA AAC CTC AAT AAT CCT CCT ATG GAA GGT ACA GGA  300
  (81) S   S   Q   S   S   K   P   S   L   N   L   N   N   P   P   M   E   G   T   G   (100)

   301 ATA TCC CTA GCA CAG TTC ATG TCT CCA AGG AGA CAG GTT ACT TCT GGA TTG GCA ACA AGG  360
 (101) I   S   L   A   Q   F   M   S   P   R   R   Q   V   T   S   G   L   A   T   R   (120)

   361 CCC AGG ATG CCA AAC AAT TCC TTT CCT CCT AAT ATT TCG ACA TTA AGC TCT CCC GTT GGC  420
 (121) P   R   M   P   N   N   S   F   P   P   N   I   S   T   L   S   S   P   V   G   (140)

   421 ATG ACA AGT AGT GCC TGT AAT AAT AAT AAC CGA TCT TAT TCA AAC ATC CCA GTA ACA TCT  480
 (141) M   T   S   S   A   C   N   N   N   N   R   S   Y   S   N   I   P   V   T   S   (160)

   481 TTA CAG GGT ATG AAT GAA GGA CCC AAT AAC TCC GTT GGC TTC TCT GCC AGT TCT CCA GTC  540
 (161) L   Q   G   M   N   E   G   P   N   N   S   V   G   F   S   A   S   S   P   V   (180)

   541 CTC AGG CAG ATG AGC TCA CAG AAT TCA CCT AGC AGA TTA AAT ATA CAA CCA GCA AAA GCT  600
 (181) L   R   Q   M   S   S   Q   N   S   P   S   R   L   N   I   Q   P   A   K   A   (200)

   601 GAG TCC AAA GAT AAC AAA GAG ATT GCC TCA ACT TTA AAT GAA ATG ATT CAA TCT GAC AAC  660
 (201) E   S   K   D   N   K   E   I   A   S   T   L   N   E   M   I   Q   S   D   N   (220)

   661 AGC TCT AGT GAT GGC AAA CCT CTG GAT TCA GGG CTT CTG CAT AAC AAT GAC AGA CTT TCA  720
 (221) S   S   S   D   G   K   P   L   D   S   G   L   L   H   N   N   D   R   L   S   (240)

   721 GAT GGA GAC AGT AAA TAC TCT CAA ACC AGT CAC AAA CTA GTG CAG CTT TTG ACA ACA ACT  780
 (241) D   G   D   S   K   Y   S   Q   T   S   H   K   L   V   Q   L   L   T   T   T   (260)

   781 GCC GAA CAG CAG TTA CGG CAT GCT GAT ATA GAC ACA AGC TGC AAA GAT GTC CTG TCT TGC  840
 (261) A   E   Q   Q   L   R   H   A   D   I   D   T   S   C   K   D   V   L   S   C   (280)

   841 ACA GGC ACT TCC AAC TCT GCC TCT GCT AAC TCT TCA GGA GGT TCT TGT CCC TCT TCT CAT  900
 (281) T   G   T   S   N   S   A   S   A   N   S   S   G   G   S   C   P   S   S   H   (300)

   901 AGC TCA TTG ACA GCA CGG CAT AAA ATT CTA CAC CGG CTC TTA CAG GAG GGT AGC CCC TCA  960
 (301) S   S   L   T   A   R   H   K   I   L   H   R   L   L   Q   E   G   S   P   S   (320)

   961 GAT ATC ACC ACT TTG TCT GTC GAG CCT GAT AAA AAG GAC AGT GCA TCT ACT TCT GTG TCA  1020
 (321) D   I   T   T   L   S   V   E   P   D   K   K   D   S   A   S   T   S   V   S   (340)

  1021 GTG ACT GGA CAG GTA CAA GGA AAC TCC AGT ATA AAA CTA GAA CTG GAT GCT TCA AAG AAA  1080
 (341) V   T   G   Q   V   Q   G   N   S   S   I   K   L   E   L   D   A   S   K   K   (360)
```

FIGURE 1

2/3

```
1081 AAA GAA TCA AAA GAC CAT CAG CTC CTA CGC TAT CTT TTA GAT AAA GAT GAG AAA GAT TTA 1140
(361) K   E   S   K   D   H   Q   L   L   R   Y   L   L   D   K   D   E   K   D   L   (380)


1141 AGA TCA ACT CCA AAC CTG AGC CTG GAT GAT GTA AAG GTG AAA GTG GAA AAG AAA GAA CAG 1200
(381) R   S   T   P   N   L   S   L   D   D   V   K   V   K   V   E   K   K   E   Q   (400)


1201 ATG GAT CCA TGT AAT ACA AAC CCA ACC CCA ATG ACG AAG GCC ACT CCT GAG GAA ATA AAA 1260
(401) M   D   P   C   N   T   N   P   T   P   M   T   K   A   T   P   E   E   I   K   (420)


1261 CTG GAG GCC CAG AGC CAG TTT ACA GCT GAC CTT GAC CAG TTT GAT CAG TTA CTG CCC ACG 1320
(421) L   E   A   Q   S   Q   F   T   A   D   L   D   Q   F   D   Q   L   L   P   T   (440)


1321 CTG GAG AAG GCA GCA CAG TTG CCA GGC TTA TGT GAG ACA GAC AGG ATG GAT GGT GCG GTC 1380
(441) L   E   K   A   A   Q   L   P   G   L   C   E   T   D   R   M   D   G   A   V   (460)


1381 ACC AGT GTA ACC ATC AAA TCG GAG ATC CTG CCA GCT TCA CTT CAG TCC GCC ACT GCC AGA 1440
(461) T   S   V   T   I   K   S   E   I   L   P   A   S   L   Q   S   A   T   A   R   (480)


1441 CCC ACT TCC AGG CTG AAT AGA TTA CCT GAG CTG GAA TTG GAA GCA ATT GAT AAC CAA TTT 1500
 81) P   T   S   R   L   N   R   L   P   E   L   E   L   E   A   I   D   N   Q   F   (500)


1501 GGA CAA CCA GGA ACA GGC GAT CAG ATT CCA TGG ACA AAT AAT ACA GTG ACA GCT ATA AAT 1560
(501) G   Q   P   G   T   G   D   Q   I   P   W   T   N   N   T   V   T   A   I   N   (520)


1561 CAG AGT AAA TCA GAA GAC CAG TGT ATT AGC TCA CAA TTA GAT GAG CTT CTC TGT CCA CCC 1620
(521) Q   S   K   S   E   D   Q   C   I   S   S   Q   L   D   E   L   L   C   P   P   (540)


1621 ACA ACA GTA GAA GGG AGA AAT GAT GAG AAG GCT CTT CTT GAA CAG CTG GTA TCC TTC CTT 1680
(541) T   T   V   E   G   R   N   D   E   K   A   L   L   E   Q   L   V   S   F   L   (560)


1681 AGT GGC AAA GAT GAA ACT GAG CTA GCT GAA CTA GAC AGA GCT CTG GGA ATT GAC AAA CTT 1740
 561) S   G   K   D   E   T   E   L   A   E   L   D   R   A   L   G   I   D   K   L   (580)


1741 GTT CAG GGG GGT GGA TTA GAT GTA TTA TCA GAG AGA TTT CCA CCA CAA CAA GCA ACG CCA 1800
 581) V   Q   G   G   G   L   D   V   L   S   E   R   F   P   P   Q   Q   A   T   P   (600)


1801 CCT TTG ATC ATG GAA GAA AGA CCC AAC CTT TAT TCC CAG CCT TAC TCT TCT CCT TTT CCT 1860
 601) P   L   I   M   E   E   R   P   N   L   Y   S   Q   P   Y   S   S   P   F   P   (620)


1861 ACT GCC AAT CTC CCT AGC CCT TTC CAA GGC ATG GTC AGG CAA AAA CCT TCA CTG GGG ACG 1920
(621) T   A   N   L   P   S   P   F   Q   G   M   V   R   Q   K   P   S   L   G   T   (640)


1921 ATG CCT GTT CAA GTA ACA CCT CCC CGA GGT GCT TTT TCA CCT GGC ATG GGC ATG CAG CCC 1980
(641) M   P   V   Q   V   T   P   P   R   G   A   F   S   P   G   M   G   M   Q   P   (660)


1981 AGG CAA ACT CTA AAC AGA CCT CCG GCT GCA CCT AAC CAG CTT CGA CTT CAA CTA CAG CAG 2040
 661) R   Q   T   L   N   R   P   P   A   A   P   N   Q   L   R   L   Q   L   Q   Q   (680)


2041 CGA TTA CAG GGA CAA CAG CAG TTG ATA CAC CAA AAT CGG CAA GCT ATC TTA AAC CAG TTT 2100
(681) R   L   Q   G   Q   Q   Q   L   I   H   Q   N   R   Q   A   I   L   N   Q   F   (700)


2101 GCA GCA ACT GCT CCT GTT GGC ATC AAT ATG AGA TCA GGC ATG CAA CAG CAA ATT ACA CCT 2160
(701) A   A   T   A   P   V   G   I   N   M   R   S   G   M   Q   Q   Q   I   T   P   (720)


2161 CAG CCA CCC CTG AAT GCT CAA ATG TTG GCA CAA CGT CAG CGG GAA CTG TAC AGT CAA CAG 2220
(721) Q   P   P   L   N   A   Q   M   L   A   Q   R   Q   R   E   L   Y   S   Q   Q   (740)
```

FIGURE 1 (con't)

3/3

```
      2221 CAC CGA CAG AGG CAG CTA ATA CAG CAG CAA AGA GCC ATG CTT ATG AGG CAG CAA AGC TTT 2280
      (741) H   R   Q   R   Q   L   I   Q   Q   Q   R   A   M   L   M   R   Q   Q   S   F   (760)


           ..................................................................
      2281 GGG AAC AAC CTC CCT CCC TCA TCT GGA CTA CCA GTT CAA ACG GGG AAC CCC CGT CTT CCT 2340
5     (761) G   N   N   L   P   P   S   S   G   L   P   V   Q   T   G   N   P   R   L   P   (780)


      2341 CAG GGT GCT CCA CAG CAA TTC CCC TAT CCA CCA AAC TAT GGT ACA AAT CCA GGA ACC CCA 2400
      (781) Q   G   A   P   Q   Q   F   P   Y   P   P   N   Y   G   T   N   P   G   T   P   (800)


      2401 CCT GCT TCT ACC AGC CCG TTT TCA CAA CTA GCA GCA AAT CCT GAA GCA TCC TTG GCC AAC 2460
10    (801) P   A   S   T   S   P   F   S   Q   L   A   A   N   P   E   A   S   L   A   N   (820)


      2461 CGC AAC AGC ATG GTG AGC AGA GGC ATG ACA GGA AAC ATA GGA GGA CAG TTT GGC ACT GGA 2520
      (821) R   N   S   M   V   S   R   G   M   T   G   N   I   G   G   Q   F   G   T   G   (840)


      2521 ATC AAT CCT CAG ATG CAG CAG AAT GTC TTC CAG TAT CCA GGA GCA GGA ATG GTT CCC CAA 2580
      (841) I   N   P   Q   M   Q   Q   N   V   F   Q   Y   P   G   A   G   M   V   P   Q   (860)


15    2581 GGT GAG GCC AAC TTT GCT CCA TCT CTA AGC CCT GGG AGC TCC ATG GTG CCG ATG CCA ATC 2640
      (861) G   E   A   N   F   A   P   S   L   S   P   G   S   S   M   V   P   M   P   I   (880)


      2641 CCT CCT CCT CAG AGT TCT CTG CTC CAG CAA ACT CCA CCT GCC TCC GGG TAT CAG TCA CCA 2700
      (881) P   P   P   Q   S   S   L   L   Q   Q   T   P   P   A   S   G   Y   Q   S   P   (900)


      2701 GAC ATG AAG GCC TGG CAG CAA GGA GCG ATA GGA AAC AAC AAT GTG TTC AGT CAA GCT GTC 2760
20    (901) D   M   K   A   W   Q   Q   G   A   I   G   N   N   N   V   F   S   Q   A   V   (920)


      2761 CAG AAC CAG CCC ACG CCT GCA CAG CCA GGA GTA TAC AAC AAC ATG AGC ATC ACC GTT TCC 2820
      (921) Q   N   Q   P   T   P   A   Q   P   G   V   Y   N   N   M   S   I   T   V   S   (940)


      2821 ATG GCA GGT GGA AAT ACG AAT GTT CAG AAC ATG AAC CCA ATG ATG GCC CAG ATG CAG ATG 2880
      941)  M   A   G   G   N   T   N   V   Q   N   M   N   P   M   M   A   Q   M   Q   M   (960)


25    2881 AGC TCT TTG CAG ATG CCA GGA ATG AAC ACT GTG TGC CCT GAG CAG ATA AAT GAT CCC GCA 2940
      (961) S   S   L   Q   M   P   G   M   N   T   V   C   P   E   Q   I   N   D   P   A   (980)


      2941 CTG AGA CAC ACA GGC CTC TAC TGC AAC CAG CTC TCA TCC ACT GAC CTT CTC AAA ACA GAA 3000
      (981) L   R   H   T   G   L   Y   C   N   Q   L   S   S   T   D   L   L   K   T   E   (1000)


      3001 GCA GAT GGA ACC CAG CAG GTG CAA CAG GTT CAG GTG TTT GCT GAC GTC CAG TGT ACA GTG 3060
30    1001) A   D   G   T   Q   Q   V   Q   Q   V   Q   V   F   A   D   V   Q   C   T   V   (1020)


      3061 AAT CTG GTA GGC GGG GAC CCT TAC CTG AAC CAG CCT GGT CCA CTG GGA ACT CAA AAG CCC 3120
      1021) N   L   V   G   G   D   P   Y   L   N   Q   P   G   P   L   G   T   Q   K   P   (1040)


      3121 ACG TCA GGA CCA CAG ACC CCC CAG GCC CAG CAG AAG AGC CTC CTT CAG CAG CTA CTG ACT 3180
      1041) T   S   G   P   Q   T   P   Q   A   Q   Q   K   S   L   L   Q   Q   L   L   T   (1060)


35    3181 GAA TAA CCACTTTTAAAGGA ATG TGA AATTTAAATAATAGACATACAGAGATATACAAATATATTATATATTTTT 3255
      (1061) E   *


      3256 CTGAGATTTTTGATATCTCAATCTGCAGCCATTCTTCAGGTCGTAGCATTTGGAGCAAAAAAAAAAAAAAAAAATCG      3332
```

FIGURE 1 (con't)